ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

161
CONTRIBUTIONS

ASSESSING IMPACTS OF CLIMATE VARIATIONS AND CHANGE ON DIFFERENT TIME SCALES

NATALIA KORHONEN

FINNISH METEOROLOGICAL INSTITUTE
CONTRIBUTIONS

No. 161

ASSESSING IMPACTS OF CLIMATE VARIATIONS AND CHANGE ON DIFFERENT TIME SCALES

Natalia Korhonen

Institute for Atmospheric and Earth System Research/Physics
Faculty of Science
University of Helsinki
Helsinki, Finland

ACADEMIC DISSERTATION in meteorology

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Auditorium E204, Physicum, Gustaf Hällströminkatu 2b, Helsinki, on March 27th 2020, at 12 noon.

Finnish Meteorological Institute
Helsinki, 2020

Title of dissertation:	Assessing Impacts of Climate Variations and Change on Different Time Scales
Author:	Natalia Korhonen Weather and Climate Change Impact Research Finnish Meteorological Institute, Finland
Thesis supervisors:	Professor Heikki Järvinen Institute for Atmospheric and Earth System Research/Physics University of Helsinki, Finland
	Docent Ari Venäläinen Weather and Climate Change Impact Research Finnish Meteorological Institute, Finland
Pre-examiners:	Professor Hans Linderholm Department of Earth Sciences University of Gothenburg, Sweden
	Professor Uwe Ulbrich Institut für Meteorologie/Klimadiagnostik und meteorologische Extremereignisse Freie Universität Berlin, Germany
Custos:	Professor Heikki Järvinen Institute for Atmospheric and Earth System Research/Physics University of Helsinki, Finland
Opponent:	Professor Markku Rummukainen Department of Physical Geography and Ecosystem Science Centre for Environmental and Climate Research (CEC) Lund University, Sweden

ISBN 978-952-336-096-9 (paperback)
ISBN 978-952-336-097-6 (pdf)
ISSN: 0782-6117

Edita Prima Oy
Helsinki 2020

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Published by Finnish Meteorological Institute (Erik Palménin aukio 1), PL 503 FIN-00101 Helsinki, Finland

Series title, number and report code of publication
Finnish Meteorological Institute Contributions 161, FMI-CONT-161
Date January 2020

Author
Natalia Korhonen

Title
Assessing Impacts of Climate Variations and Change on Different Time Scales

Climate change refers to a change in the mean state of the climate that persists for an extended period, typically 30 years or longer. The natural inter-annual variability of climate refers to internal variation of the climate system in shorter time-scales. In this thesis I have studied the climate of the last glacial and its impact on human population sizes in Europe during the end of the last glaciation, the change in forest fire danger and strong winds in Europe under the on-going human-induced climate change, and the relationship between the stratospheric winds and the phase of the Arctic Oscillation in present climate.

A regression model was developed to downscale low-resolution dynamical EMIC simulations to regional scale. The regression model was calibrated by gridded data of regional scale resolution observations of present day climate and simulations of glacial climate. The downscaled climate was used in estimating the size of human population in Europe during the end of the last glaciation, between 30,000 and 13,000 years ago. The simulated changes in human population size correlated significantly with an independent archeological data of changes in human population size.

The change in the forest fire danger in Europe was investigated by ERA-Interim and ERA-40 reanalysis. The forest fire danger was found to have increased in Southern and Eastern Europe during the period 1980–2012, whilst no significant trend was found elsewhere in Europe.

The projected changes in the geostrophic wind speeds under human-induced climate change in Northern Europe during the current century were explored from simulations of nine general circulation models. According to the simulations, the changes in mean and extreme wind speeds are going to be small; in parts of northwestern Russia and southern Baltic Sea the winds might increase by 2-4% and over the Norwegian Sea the winds might decrease by 2-8%.

In this thesis the connection between the stratospheric winds and surface Arctic Oscillation was studied statistically. The found stratospheric connection was applied in post-processing the European Centre for Medium-Range Weather Forecasts two-week mean temperature reforecasts for weeks 3–4 and weeks 5–6 in Northern Europe during boreal winter, and the skill scores of those weeks were slightly improved.

Publishing unit
Finnish Meteorological Institute, Weather and Climate Change Impact Research

Classification (UDC)
551.509.314, 551.509.33, 551.510.529, 551.583.3

Keywords
climate change, glacial climate, strong winds, stratosphere-troposphere interaction, forest fire risk, sub-seasonal forecasts

ISSN and series title
ISSN 0782-6117 Finnish Meteorological Institute Contributions

ISBN
ISBN 978-952-336-096-9 (paperback)
ISBN 978-952-336-097-6 (pdf)

DOI
https://doi.org/10.35614/isbn.9789523360976

Language
English

Pages
136

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Julkaisija
Ilmatieteen laitos
(Erik Palménin aukio 1), PL 503
00101 Helsinki, Finland

Julkaisun sarja, numero ja raporttikoodi
Finnish Meteorological Institute Contributions 161,
FMI-CONT-161
Päiväys Tammikuu 2020

Tekijä
Natalia Korhonen

Nimeke
Ilmaston muutosten ja vaihteluiden vaikutusarviointeja eri aikaskaaloissa

Ilmaston muutoksella tarkoitetaan ilmaston keskimääräisen tilan muutosta pitkän ajan, yleensä 30 vuoden tai sitä pidemmän ajanjakson aikana. Ilmaston luonnollinen vuosien välinen vaihtelu taas kuvaa lyhemmässä aikaskaalassa ilmastojärjestelmän sisäistä vaihtelua. Tässä väitöskirjassa on tutkittu viime jääkauden lopun luonnollisen ilmaston muutoksen vaikutusta ihmisten määrään Euroopassa, ihmisten aiheuttaman ilmastonmuutoksen vaikutusta metsäpalovaaran ja voimakkaisiin tuuliin Euroopassa sekä stratosfäärien tuulien ja Arktisen värähtelyn tilastollista yhteyttä nykyilmastossa.

Tässä väitöskirjassa karkean erottelukyvyn jääkausi-ilmastosimulaatio alueellistettiin regressiomallilla, joka hyödynsi hilamuotoista aineistoa sekä nykyilmastosta että yleisen kiertoliikkeen mallin jääkausisimulaatioista. Mallinnettujen väestömäärien vaihtelut korreloivat merkitsevästi riippumattoman arkeologisen aineiston väestömäärän muutosten arvioiden kanssa.

Ilmastollisten metsäpaloherkkyyteen vaikuttavien suureiden viimeaikaista trendiä Euroopan alueella tutkittiin ERA-Interim- ja ERA-40-uusanalyysiaineistojen avulla. Metsäpaloriskin havaittiin kasvaneen Etelä- ja Itä-Euroopassa jaksolla 1980–2012, kun taas muualla Euroopassa selvää trendiä ei ollut havaittavissa.

Ihmisten aiheuttaman ilmaston muutoksen vaikutusta tuulisuuteen Pohjois-Euroopassa kuluvan vuosisadan aikana tutkittiin yhdeksällä yleisen kiertoliikkeen mallilla. Tulosten perusteella keskimääräisten ja voimakkaimpien geostrofisten tuulien muutokset ovat muutaman prosentin luokkaa; mallinnukset kuluvalle vuosisadalle näyttivät näiden tuulien voimistuvan joitakin prosentteja osassa Luoteis-Venäjää ja Itämeren eteläosia ja heikkenevän muutamia prosentteja Norjanmerellä.

Tässä väitöskirjassa tutkittiin stratosfäärin tuulien ja arktisen värähtelyn yhteyttä tilastollisesti. Hyödyntämällä stratosfääristä saatavaa signaalia pystyttiin 3-6 viikon päähän ulottuvia säämalleilla tehtyjä Pohjois-Euroopan viikkokeskilämpötilojen ennusteita hieman parantamaan.

Julkaisijayksikkö
Ilmatieteen laitos, Sään ja ilmastonmuutoksen vaikutustutkimus

Luokitus (UDK)
551.509.314, 551.509.33, 551.510.529, 551.583.3

Asiasanat
ilmaston muutos, jääkauden ilmasto, kova tuuli, stratosfääri-troposfääri-vuorovaikutus, metsäpaloriski, pitkät sääennusteet

ISSN ja avainnimeke
ISSN 0782-6117 Finnish Meteorological Institute Contributions

ISBN
ISBN 978-952-336-096-9 (nid.)
ISBN 978-952-336-097-6 (pdf)

DOI
https://doi.org/10.35614/isbn.9789523360976

Kieli
englanti

Sivumäärä
136

ACKNOWLEDGEMENTS

I thank my supervisors Professor Heikki Järvinen and Docent Ari Venäläinen for their guidance and support. I thank all my co-authors for their contributions, especially Adjunct professor Hilppa Gregow and Dr. Otto Hyvärinen. I thank all my co-workers at our Weather and Climate Change Impact Research unit (SIV), especially Antti Mäkelä, Pentti Pirinen, Tiina Ervasti, Terhi Laurila and Anna Luomaranta for their support.

I am grateful to Professor Markku Rummukainen to be as opponent for my thesis, Professor Heikki Järvinen to serve as a custos and Professors Uwe Ulbrich and Hans Linderholm for pre-examining this thesis.

I thank all the teachers in my life, especially my parents and late grandparents, my first class teacher Ilkka Vesala, who taught me to read; my first physics teacher Liisa Holopainen; and my physics teacher in senior high school, Dr. Heikki Saari, who taugh us pupils that "in physics you can never be sure". I thank my lecturers at the University of Helsinki for impressive lectures.

I want to thank the Finnish Meteorological Institute for providing working facilities. This work was funded by the Academy of Finland, Posiva Oy, and the Finnish Meteorological Institute.

Finally, my friends and family, thank you for being there with me, especially my husband Janne, and the lion and the sun I have given birth to: Leo and Ella.

Natalia Korhonen
Helsinki, January 2020

CONTENTS

ABBREVIATIONS

AD	Anno Domini
AO	Arctic Oscillation
BP	Before Present
CCSM3	Community Climate System Model 3
CLIMBER-2	CLIMate and BiosphERe model 2
CMIP3	Coupled Model Intergovernmental Project 3
CMIP5	Coupled Model Intergovernmental Project 5
CO2	Carbon dioxide
CRPS	Continuously Ranked Probability Score
CRPSS	Continuously Ranked Probability Skill Score
ECMWF	European Centre for Medium-Range Weather Forecasts
EMIC	Earth system Model of Intermediate Complexity
EPICA	European Project for Ice Coring in Antarctica
EQBO	easterly QBO
ERA	ECMWF re-analysis
ERF	Extended Range Forecasts
FWI	Canadian Fire Weather Index
GAM	Generalized Additive Model
GCM	General Circulation Model
GEV	Generalized Extreme Values
IPCC	Intergovernmental Panel on Climate Change
kyr	1,000 years
LGM	Last Glacial Maximum
QBO	Quasi-Biennial Oscillation
RCA3	Rossby-Centre regional Atmospheric climate model 3
RCP	Representative Concentration Pathway
SICOPOLIS	SImulation COde for POLythermal Ice Sheets
SRES	Special Report on Emission Scenarios
SWI	Stratospheric Wind Indicator
T2m	surface air temperature
WQBO	westerly QBO
ZMZW	Zonal Mean Zonal Wind speed

LIST OF ORIGINAL PUBLICATIONS

I Gregow, H., Ruosteenoja, K., **Pimenoff, N.** and Jylhä, K. 2011. Changes in the mean and extreme geostrophic wind speeds in Northern Europe until 2100 based on nine global climate models. *International Journal of Climatology*, 32, 1834–1846. doi: 10.1002/joc.2398

II Venäläinen, A., **Korhonen, N.**, Hyvärinen, O., Koutsias, N., Xystrakis, F., Urbieta, I. R, and Moreno, J. M. 2014. Temporal variations and change in forest fire danger in Europe for 1960–2012. *Natural Hazards and Earth System Sciences*, 14, 1477–1490. doi: 10.5194/nhess-14-1477-2014

III **Korhonen, N.**, Venäläinen, A., Seppä, H., and Järvinen, H. 2014. Statistical downscaling of a climate simulation of the last glacial cycle: temperature and precipitation over Northern Europe. *Climate of the Past*, 10, 1489–1500. doi:10.5194/cp-10-1489-2014

IV Tallavaara, M., Luoto, M., **Korhonen, N.**, Järvinen, H. & Seppä, H. 2015. Human population dynamics in Europe over the Last Glacial Maximum. *Proceedings of the National Academy of Sciences of the United States of America*, 112: 8232–8237. https://doi.org/10.1073/pnas.1503784112

V **Korhonen, N.**, Hyvärinen, O., Kämäräinen, M., Richardson, D. S., Järvinen, H., and Gregow, H. 2019. Adding value to Extended-range Forecasts in Northern Europe by Statistical Post-processing Using Stratospheric Observations, *Atmos. Chem. Phys. Discuss.*, https://doi.org/10.5194/acp-2019-679, in review.

Author's contribution

Natalia Korhonen was solely responsible for the summary of this thesis. In **Paper I** Natalia Korhonen (née Pimenoff) performed the gridded extreme value calculations and participated in writing the gridded extreme value calculations related matters. In **Paper II** Natalia Korhonen participated in the calculation and writing of the Canadian Fire Weather Indexes. In **Paper III** and in **Paper V** Natalia Korhonen had the main responsibility for the planning, the data analysis and writing of the papers. In **Paper IV** Natalia Korhonen performed the statistical downscaling of the model-based climate data and participated in writing the paper.

Paper I and **Paper IV** have been in the theses of Hilppa Gregow and Miikka Tallavaara, respectively.

1. INTRODUCTION

The global climate is changing continuously. The causes and the time-scales of the climate changes vary (Harvey, 2000). The climate changes and variations can be caused either by external climate forcings or by the internal dynamics of the climate system. External climate forcings include natural phenomena such as continental drifts (Wessel & Müller 2015), variations in solar insolation (Berger and Loutre 1991) and solar irradiance (Haigh 1996; Ammann et al. 2007), volcanic eruptions (Ammann et al. 2007, Anchukaitis et al. 2010), as well as human-induced changes in the composition of the Earth's atmosphere (IPCC 2013). The climate system's own internal dynamics comprise of interactions and feedback effects between the various components of the climate system, e.g., atmosphere-biosphere interaction (Frank et al. 2010), the ice-albedo feedback effect (e.g., Joughin and Alley 2011), and troposphere-stratosphere interaction (Manzini et al. 2014). The climate changes might be slow – like those driven by continental drifts where the slowest changes take hundreds of millions of years - or fast like cooling caused by large volcanic eruptions. The climate also variates from year-to-year (inter-annually) due to large-scale atmospheric circulation variability, driven among other things by variation in ocean surface temperatures (Alexander et al. 2009) and ice and snow cover (Cohen et al. 2014).

Climate changes have versatile impacts. The general aim of this thesis has been to increase our understanding on a set of past and future impacts of climate change (at regional scale) over Europe. For assessing climate variation and change and their impacts, we need appropriate tools. The complex interactions of the climate system are simulated by numerical models. These climate models simulate the interactions quantitatively by differential equations based on the basic laws of physics, fluid motion and chemistry. In this work I demonstrate how the state of the climate system can be simulated in time scales ranging from weeks to 100,000 years. In addition to this, I introduce examples of climate change impact studies.

I start with investigating the climate of the last glacial cycle (126,000 to 0 years ago) and its impact on human population size in Europe during the end of the last glaciation, about 30,000 to 13,000 years ago. Humans inhabited Europe (e.g., Vermeersch, 2005) even during the coldest stage of the last glaciation, about 21,000 years ago, when the global climate was about 4 to 7 degrees cooler than at present (Masson-Delmotte et al. 2013). It is, however, still unclear how much the climate variations during the end of the last glacial affected human population size in Europe during that time. Hence, in this thesis I study:

1. how to reproduce spatial patterns of the climate in Europe during the last glacial cycle, 126,000 to 0 years ago (**Paper III**), and

2. whether climate was an important driver of the hunter-gatherer population dynamics in Europe during the end of the last glaciation, 30,000 to 13,000 years ago **(Paper IV**).

As natural climate changes have caused significant ecosystem shifts and species extinctions during the past million years (IPCC, 2014), present day species and ecosystems of the Earth are affected by the ongoing human-induced climate change (IPCC, 2018). As forest ecosystems have an important role in climate change mitigation as carbon sinks, and in economy as "the green gold", their wellbeing is important. To support planning the forest management practices, estimates of

the climate change impacts on forest ecosystems under human-induced climate change are needed. In this thesis I investigate the human-induced climate changes effect on windiness and forest fire risk, more precisely:

3. how the climate-based forest fire risk has changed in Europe between 1960 and 2012 under the on-going human-induced climate change (**Paper II**), and

4. how the mean and extreme winds are projected to change during the 21st century over Northern Europe in projections performed with nine global circulation models (GCMs) employing the Special Report on Emission Scenarios (SRES, Nakicénovic et al., 2000) A1B, A2 and B1 scenarios (**Paper I**).

According to Coupled Model Intergovernmental Project 5 (CMIP5) model simulations, the human-induced climate change might lead also to changes in the stratospheric flow, for example the increase of greenhouse gases in the atmosphere might have a decreasing impact on the strength of the winter time stratospheric polar vortex (Manzini et al. 2014). The stratospheric polar vortex has significant impact on the winter time surface weather in Northern Europe (Kidston et al. 2015), and as there is a potential these stratospheric conditions might change in the course of human-induced climate change, it is important to know their connection to the surface weather.

During boreal winter the stratospheric polar vortex influences the surface weather in the Northern Hemisphere within weeks or months (Baldwin and Dunkerton 1999, Limpasuvan et al. 2005, Thompson et al. 2002, Tomassini et al. 2012, Kidston et al. 2015): the weaker (stronger) than average stratospheric polar vortex is connected to colder (warmer) than average surface temperatures in Northern Europe. Further, the stratospheric polar vortex is influenced by the Quasi-Biennial Oscillation (QBO), a quasiperiodic oscillation of the equatorial zonal wind between downwards propagating easterlies and westerlies in the tropical stratosphere with a mean period of 28 to 29 months (Baldwin et al., 2001); easterly (westerly) QBO often coincides with weaker (stronger) stratospheric polar vortex (Holton and Tan, 1980). The QBO–stratospheric polar vortex–surface weather connection holds a potential to improve Extended-range forecasts (ERF; lead time 14 to 46 days, Domeisen 2019a, 2019b), which forecasting skills of the surface weather are still rather modest in the Northern latitudes. In this thesis my aim is to utilize the connection between the stratospheric polar vortex and the Northern Europe's winter surface weather in improving ERFs, more precisely I study:

5. how the teleconnection between the QBO and the stratospheric polar vortex between 1981-2016 could be utilized in improving extended range temperature forecasts in Northern Europe (**Paper V**).

2. BACKGROUND

In the following Sections, we briefly introduce needful background knowledge about the main climate variations and changes studied in this thesis.

2.1 Glacial climate

The continental drifts (Wessel & Müller 2015) alter the global climate by altering the Earth's land-ocean distribution, the ocean currents, and land form, and also by affecting the composition of the atmosphere. The continental drifts are considered to be the fundamental reason for the glacial times that have occurred on the time-scales of hundreds of millions of years.

Within glacial times the climate has varied between glacials and interglacials driven largely by changes in Earth's orbit (Crucifix et al. 2006; Berger and Loutre 1991; Masson-Delmotte et al. 2013). According to paleoclimatic reconstructions from the ice cores and ocean sediments, the Earth's climate has during the last 650,000 years varied between glacial and interglacial conditions with a strong periodicity of approximately 100,000 years (EPICA community members 2004). The orbital theory states that glaciations are triggered by minima in summer insolation in the Northern high latitudes. This enables the snow aggregated during the winter to stay over the summer and therefore to gradually accumulate, generating the Northern Hemisphere continental ice sheets. Figure 1a) depicts the solar insolation in June at 60 °N between 40,000 years Before Present (BP, i.e., before 1 January 1950 AD) and 2050 AD calculated by Berger and Loutre (1991).

The last glaciation, the Weichselian glaciation began about 115,000 years ago and ended about 11,500 years ago. Figure 1c shows a proxy of local temperature in the Antarctic between 40 kyr BP and 0.1 kyr BP. During the coldest stage of the Weichselian glacial, the Last Glacial Maximum (LGM) about 20 kyr BP, the global mean temperature is estimated to have been 3 to 8 °C lower than the pre-industrial climate (Masson-Delmotte et al. 2013) and large continental ice sheets covered Northern Europe and Northern North America (Ehlers and Gibbard, 2004). The well-known insolation variations do not alone explain these variations. They have indeed been amplified by the interactions and feedbacks of the Earth's climate system. For example, if the summer insolation in the Northern high latitudes decreases due to orbital changes, the climate cools and increasing amounts of precipitation are achieved as snow. This leads to whiter surfaces which reflect more of the incoming solar radiation causing more cooling, which causes more snow and ice, and so on, in a self-reinforcing cycle. This is called the 'ice-albedo feedback' and it works the opposite in warming climate; warmer temperatures melt snow and ice, which reveals darker surfaces that have previously been beneath the snow and ice, and these darker surfaces absorb more of the solar radiation, leading to more warming.

Other important feedbacks of the climate system include the responses of the carbon cycle, the hydrological cycle and the terrestrial biosphere (Crucifix et al. 2006). Proxy data from ice cores reveal that during the glacial cycles the atmospheric CO_2 concentration has varied between 180 ppm and 300 ppm (EPICA community members 2004; Petit et al. 1999). Figure 1b) depicts the atmospheric CO_2 concentration between 40 kyr BP and 0.1 kyr BP reconstructed from ice cores. During the LGM the global cooling resulted from the presence of the Northern Hemisphere ice sheets, the decreased atmospheric CO_2 concentration, shrunken vegetation cover and increased atmospheric dust content (e.g., Otto-Bliesner et al. 2006, Shaffer & Lambert 2018).

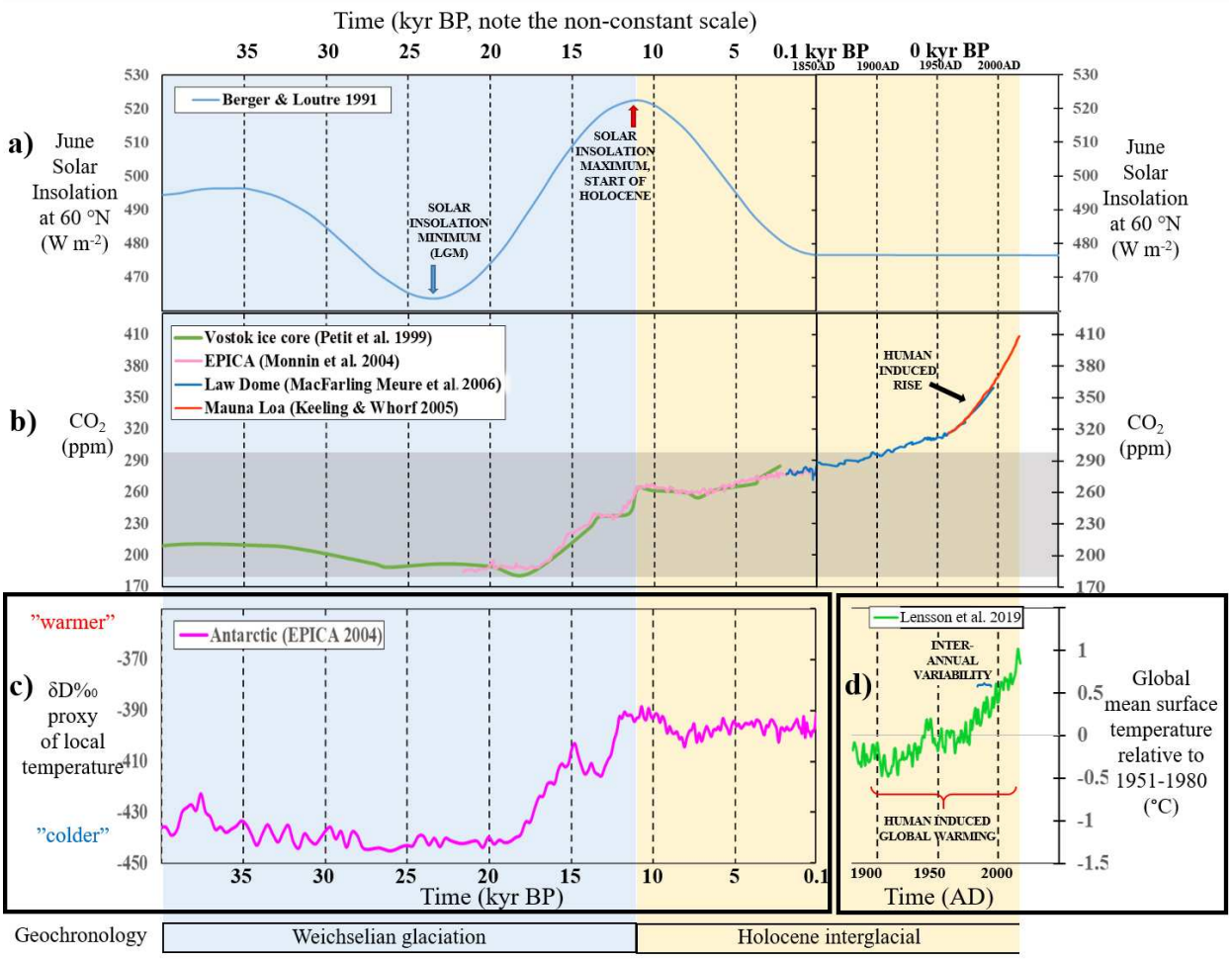

Figure 1 a) June solar insolation at 60 °N between 40 kyr BP and 2050 AD (Berger and Loutre, 1991). b) Reconstructed (Vostok ice core, EPICA and Law Dome) and direct measurements (Mauna Loa) of the atmospheric CO_2 concentration between 40 kyr BP and 2018 AD. The grey shading shows the range of the reconstructed atmospheric CO_2 concentration during 650,000 years before the industrialization (EPICA community members 2004). c) δD%0, a proxy of local temperature at the Antarctic between 40 kyr BP and 0.1 kyr BP (EPICA community members 2004). d) Global land-ocean temperature index between 1880 AD and 2018 AD (Lenssen et al. 2019). The blue (yellow) shading indicates the Weichselian glaciation (Holocene interglacial). Figure compiled by the author from the mentioned sources.

2.2 Human-induced climate change

After the solar insolation maximum about 11.5 kyr BP as depicted in Figure 1a, the Earth's climate warmed so much that the Weichselian glacial ended and the current interglacial, the Holocene, started. During the Holocene interglacial, the atmospheric carbon dioxide (CO_2) concentration was stable between 260 and 280 ppm for more than 10,000 years (Figure 1b, Monnin et al. 2004). During the industrialized era (the last 270 years), human activities have increased the atmospheric CO_2 concentration from its pre-industrial values of 280 ppm to about 418 ppm (MacFarling Meure et al. 2006; Keeling and Whorf 2005, see Figure 1b), primarily by the combustion of fossil fuels

and deforestation, but also by cement production and other land-use changes. While human activities affect the climate change in many direct and indirect ways, the CO_2 emissions are considered the single largest anthropogenic factor contributing to the observed global mean temperature rise of ca. 1 °C during the last 150 years (IPCC, 2018), depicted in Figure 1d. Depending on the greenhouse gas emissions the global mean temperature is projected to rise 0.5…4 degrees during the ongoing 21st century, higher emissions leading to higher changes in the mean temperature (IPCC, 2018; Figure 2).

2.3 Inter-annual variability of winter temperatures in Northern Europe

The see-saw of the annual mean global mean temperature (Lenssen et al. 2019) in Figure 1d depicts the inter-annual variability of the global mean temperature. It is different from the global mean temperature change which is identified by changes in the mean temperature that persists for an extended period, typically decades or longer. Inter-annual variability of climate is caused by the redistribution and changes in the amount of energy around the globe leading to changes in climate variables such as pressure and temperature. The amount and movement of energy are driven among other things by variation in ocean surface temperatures, volcanic activity, and ice and snow cover.

The inter-annual variability of climate is particularly strong in the high latitudes, e.g., in Finland the annual mean temperature was in 1981-2010 on average about 2.3 °C with the minimum of -0.4 °C in year 1985 and the maximum of 3.8 °C in year 1989 (Mikkonen et al. 2015). In Northern Europe one of the important indicators of the large-scale weather patterns is the phase of the Arctic Oscillation (AO). The AO is a climate pattern characterized by winds circulating counterclockwise around the Arctic at around 55°N latitude (Baldwin and Dunkerton 1999). The positive AO phase, as schematically depicted in Figure 3a, is characterized by lower than average pressure in the Arctic and strong westerly winds around the North pole keeping the cold Arctic air locked in the polar region and bringing milder and wetter than average weather to Northern Europe. The negative AO phase, as schematically depicted in Figure 3b, is characterized by higher than average surface pressure in the Arctic, and the meandering and/or weakening of the polar jet stream and tropospheric jet stream enabling cold arctic/polar air outbreaks to Northern Europe.

The difference in annual mean temperatures emphasize the effect of the dominating large-scale weather patterns: in 1985 negative AO indexes dominated during winter months and the annual mean temperature in Finland was -0.4 °C, whereas in 1989 positive AO indexes dominated during most of the whole year and the annual mean temperature in Finland was 3.8 °C (Mikkonen et al. 2015).

During winter time the phase of the AO is affected by the northern stratospheric polar vortex. The stratospheric polar vortex is an upper-level low-pressure area that forms over both the northern and southern poles during winter due to the growing temperature gradient between the pole and the tropics. Strong westerly winds circulate the stratospheric polar vortex, isolating the gradually cooling polar cap air. The strength of the northern stratospheric polar vortex varies from year to year and can be indicated by, e.g., the zonal mean zonal wind speed (ZMZW) at 60 °N and 10 hPa or polar cap temperatures. The stronger the circumpolar winds and the colder the polar cap temperatures are, the stronger is the polar vortex. Planetary waves from the troposphere disturb the northern stratospheric polar vortex, leading to meandering and weakening of the westerlies and

occasionally to reverse, i.e., easterly flow (Schoeberl, 1978). This weakening of the stratospheric polar vortex also leads to warming of the polar cap temperatures, sometimes even > 30–40 °C within several days. A warming of this magnitude together with a reversal of the ZMZW at 10 hPa at 60 °N is commonly defined as a major sudden stratospheric warming (SSW), albeit other definitions also occur (Butler et al. 2015). During boreal winters, the strength of the stratospheric polar vortex influences the surface weather in the Northern Hemisphere within weeks or months (Baldwin and Dunkerton 1999, Limpasuvan et al. 2005, Thompson et al. 2002, Tomassini et al. 2012, Kidston et al. 2015): the weaker (stronger) than average stratospheric polar vortex is connected to negative (positive) Arctic Oscillation (AO) and colder (warmer) than average surface temperatures in Northern Europe.

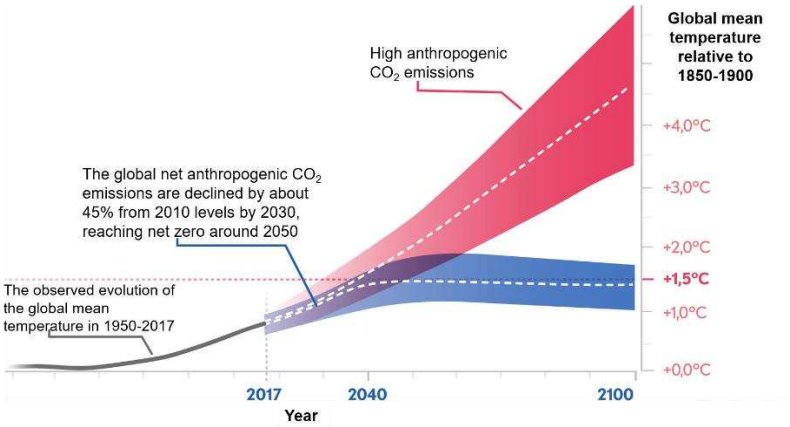

Figure 2. *The ongoing global warming and the projected rise in the global mean temperature during the current century under different greenhouse gas emission scenarios (Figure source: IPCC (2018), Finnish Meteorological Institute, Finland's Ministry of Environment, Ilmasto-opas.fi).*

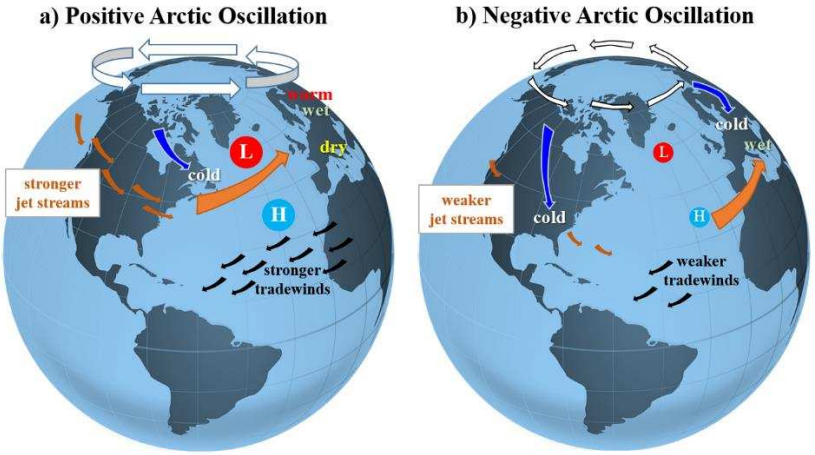

Figure 3 *A Schematic figure of the positive AO (a) and negative AO (b) and their effects. The red encircled L (blue encircled H) represent centers of low (high) pressure systems over the North Atlantic. Figure by the author.*

During boreal winter the quasi-biennial oscillation (QBO), a quasiperiodic oscillation of the equatorial zonal wind between downwards propagating easterlies and westerlies in the tropical stratosphere with a mean period of 28 to 29 months (Baldwin et al. 2001), affects the polar vortex. The easterly (westerly) QBO often coincides with weaker (stronger) than average polar vortex. There is no precise consensus of the mechanisms of this tropical-extratropical connection, but the most common explanation is that the QBO affects the polar vortex via the Holton Tan effect: During easterly QBO, small amplitude planetary waves are reflected back towards the North Pole weakening the polar vortex (Holton and Tan 1980, 1982; Watson and Gray 2014, Gray et al. 2018). In model simulations by Garfinkel et al. (2018) a weakened stratospheric polar vortex during the easterly QBO phase compared to the westerly phase was found in early winter (October–December).

2.4. Simulation of climate in different time scales

Past, present and future climates on time-scales of 100 years are commonly studied by atmospheric General Circulation Models (GCMs) coupled with modules simulating the marine, land biospheres and the sea ice. The resolution of such global GCMs is usually 100–300 km. The human-induced climate change is projected with GCMs. In these models the effects of changes in the forcing conditions – e.g. enhanced greenhouse gas concentrations – to the climate system are simulated. There are large uncertainties in the climate change projections, uncertainties like natural variation of climate, the uncertainties in the future emissions and model's simplicities. Moreover, when projecting the feedback effects of the ecosystems, the uncertainty is caused also by the uncertainty of human action and nature's ability to adapt to the changing environment. The GCMs project that the Earth's climate is going to warm further during the next centuries (IPCC 2013), as the anthropogenic emissions continue to rise the atmospheric CO_2 concentration. Further, the global mean temperature is projected to rise 0.5…4 degrees during the ongoing 21st century, depending on the greenhouse gas emissions (IPCC, 2018; Figure 2).

GCMs with coupled ice sheets are computationally too expensive to simulate full glacial cycles (100,000 years). Therefore, the GCMs have only been used to simulate glacial climate with prescribed ice sheets. For example Otto-Bliesner et al. (2006); Brandefelt and Otto-Bliesner (2009), and Brady et al. (2013) have done time-slice simulations of the distant past climate, and Smith and Gregory (2012) have simulated the whole last glacial cycle by atmosphere-ocean general circulation models with prescribed ice sheets.

Computationally suitable models for modelling glacial climate in time-scales of 100,000 years are Earth system models of intermediate complexity (EMIC; Claussen et al., 2002; Petoukhov et al., 2005). In terms of their complexity the EMICs lie between simple energy-balance models and comprehensive general circulation models (GCMs). The spatial resolution of the EMICs is low, which enables very long simulations with reasonable computing time. The atmospheric, ocean, sea ice, land surface, terrestrial vegetation and ice sheet components of the EMICs are simplified further decreasing the computing costs. The EMICs are run by prescribed solar forcing (insolation), and atmospheric greenhouse gas concentrations. EMICs are suitable for simulating climate in a time scale of 100,000 years, however due to their coarse resolution, the output is not directly usable for regional scale climate risk assessments or bioclimatic studies.

3. MATERIALS AND METHODS

In the following sections we introduce the data and methods used in Papers I to V.

3.1 Glacial climate in Europe

In **Paper III** we downscaled the coarse output an EMIC simulation of the last glacial cycle (126 kyr BP to 0 kyr BP) into regional scale, and in **Paper IV** we used the downscaled climate data representing 30 to 13 kyr BP in estimating human population size in Europe during that period. Figure 4 depicts the flow chart of the materials and topics of **Paper III** and **Paper IV** and Table 1 summarizes the data used in Papers III and IV.

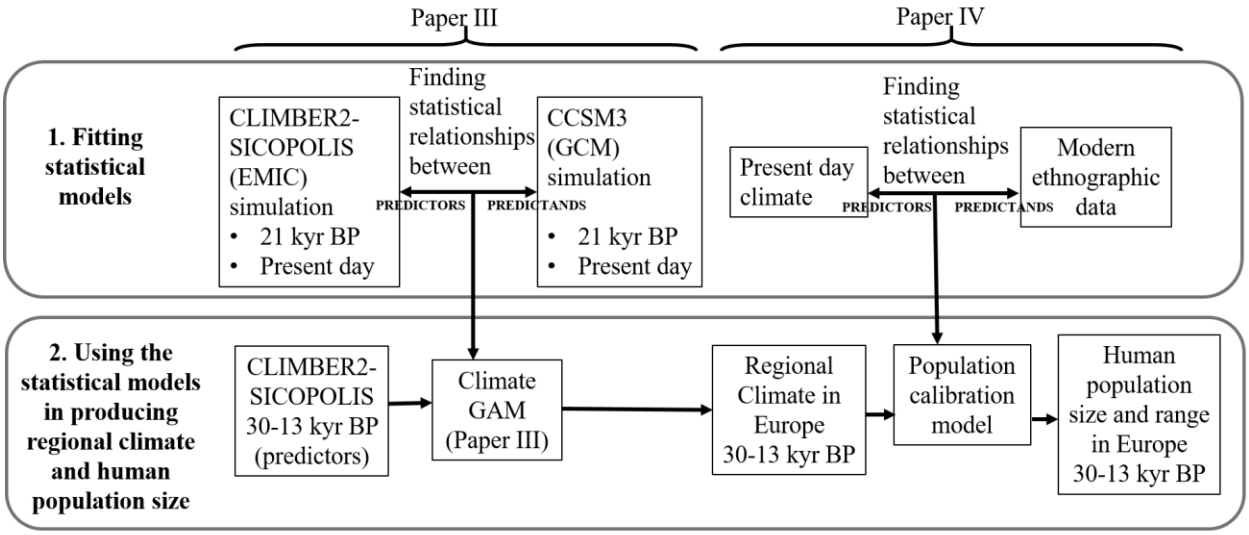

Figure 4 Flow chart showing the material and topics of Papers III and IV. Figure by the author.

3.1.1. Downscaling of a glacial climate simulation (Paper III)

Glacial climate was simulated by the EMIC CLIMBER-2 (Ganopolski et al. 2010; Petoukhov et al. 2000; Ganopolski et al. 2001). The CLIMBER-2 comprises of six earth system components: atmosphere, ocean, sea ice, land surface, terrestrial vegetation and ice sheets. The first five components are represented by coarse-resolution modules of intermediate complexity. The ice-sheet component is a three-dimensional polythermal ice-sheet model SICOPOLIS (Greve, 1997; Calov et al. 2005).

To downscale the coarse output of the CLIMBER-2 into suitable regional scale, we used GAMs. In GAMs (Hastie and Tibshirani, 1990; Wood, 2006) the statistical expectation (E) of the predictand (Y) were modelled using the sum of univariate spline functions of the p predictors ($X1$,..., Xp), such that

$$E\left(Y\middle|X_1 \dots X_p\right) = \sum_{j=1}^{p} fj(Xj), \quad (1)$$

where the potentially non-linear spline functions fj had a non-parametric form.

The calibration of GAMs is data driven, and the shape of the response is not forced to any parametric form. GAMs are called semiparametric models, as the probability distribution of the predictand should be known. The downscaled parameters, i.e., the predictands, were annual mean precipitation and annual mean temperature (see Table 1). As temperature data classically satisfy the normality assumption, the response variable mean temperature was expected to follow a normal distribution. According to Cheng and Qi (2002), the cumulative precipitation data can be modelled by a lognormal distribution, hence the total precipitation response variable was log-transformed and expected to follow a normal distribution. For fitting the GAM, all the data were to be represented at the same spatial resolution. We used a resolution of 1.5° in latitudinal and 0.75° in longitudinal dimensions. The GAMs were fitted by an R package "mgcv", for Europe (36–70° N, 10° W–69° E). The GAMs were calibrated by observations of present day climate (Mitchell and Jones, 2005) and by the output of CCSM3 GCM simulation of LGM climate (Brady et al. 2013).

The best fit for downscaling the annual mean precipitation were attained with following predictors: CLIMBER-2 bi-linearly interpolated annual mean precipitation, latitude, longitude, elevation of the grid point, and direction of the steepest slope. Moreover, the best fit to downscale annual mean temperature were attained by following predictors: CLIMBER-2 bi-linearly interpolated temperature, latitude, longitude, elevation and the shortest distance to the ice sheet margin.

The downscaled annual mean temperatures were compared to two pollen-based reconstructions of annual mean temperature by Heikkilä and Seppä (2003) and Antonsson et al. (2006). Moreover, the downscaled annual mean surface temperature and precipitation of the 44 kyr BP climate were compared to RCA3 regional climate model simulation output by Kjellström et al. (2010) representing 44 kyr BP annual mean surface temperature and precipitation over Europe.

3.1.2 Human population calibration model (Paper IV)

Climate envelope models use associations between climate and the occurrences of species to estimate how populations change in response to climate change (Pearson and Dawson 2003). The human population size in Europe between 30 and 13 kyr ago was simulated by a population calibration model constructed by ethnographic data on modern terrestrially adapted mobile hunter-gatherers and their climatic space (Binford 2001), see Table 1. This population calibration model used potential evapotranspiration, water balance and the mean surface temperature of the coldest month as input to predict the hunter-gatherer presence and population density. The climatic input were achieved by statistical downscaling (section 3.1.1) of the CLIMBER-2 climate model simulation of the last glacial cycle (Ganopolski et al. 2010). The simulated range and size of the human population was compared to archeological data (Vermeersch 2005), see Table 1.

Table 1 Data used in Papers III–IV.

Paper	Data	Source	Time period	Area
III	Monthly mean: • surface temperature • total precipitation • direction of the steepest slope • elevation • the shortest distance to the ice sheet margin	CLIMBER-2-SICOPOLIS (Ganopolski et al. 2010)	126–0 kyr BP	Europe
		GCM: CCSM3 (Brady et al. 2013)	21 kyr BP 0 kyr BP	Europe
		RCM: RCA3 (Kjellström et al. 2010)	44 kyr BP	Europe
	Monthly mean: • surface temperature • total precipitation	Mitchell and Jones (2005)	1961-1990	Europe
	• pollen-based reconstructions of annual mean temperature	Heikkilä and Seppä (2003) Antonsson et al. (2006)	10–0 kyr BP	Laihalampi (Finland) Gilltjärnen (Sweden)
IV	• ethnographic data from 127 hunter-gatherer populations and their climatic space	Binford (2001)	Recent historical times	Global
	• potential evapotranspiration • water balance • mean surface temperature of the coldest month	World CLimDatabase (Hijmans et al., 2005)	Means of 1950–2000	Global
		CLIMBER-2-SICOPOLIS (Paper III)	30–13 kyr BP	Europe
	• archaeological proxy of population size based on radiocarbon dates	Vermeersch (2005)	30–13 kyr BP	Europe

3.2 Human-induced climate change impacts in Europe

3.2.1 Trend analysis of the Canadian Fire Weather Index (FWI) (Paper II)

We studied the change in the forest fire danger in specified land areas of Northern, Western, Eastern and Southern Europe (see Figure 11 (a) and (c) for the specified land areas) in 1980-2012 by ERA-Interim reanalysis (Dee et al. 2011), and in 1960-1999 by ERA-40 reanalysis (Uppala et al. 2005), see Table 2. The spatial resolution was 2.5° × 2.5° for ERA-40 and 1.5° × 1.5° for ERA-Interim. We used a much used measure for forest fire danger: the Canadian Fire Weather Index (FWI, Van Wagner 1987). The FWI was calculated using the precipitation sum of the previous 24 h, midday temperature, relative humidity, and wind speed. We used the R package "fume" (Bedia et al. 2015) to calculate the March-September mean of the FWI and their trend, see Figure 5. The trend of the March-September mean FWI was analyzed by the Mann-Kendall test (Mann, 1945; Kendall, 1975).

Further, we investigated how the number of days with high fire danger risk as assessed using FWI have changed within 1960-2012. The threshold of FWI values larger than 20 was regarded as applicable for high fire danger risk for cool climates and was used in whole Europe and in the specified areas. For southern Europe we used the threshold of FWI values larger than 45 for high fire danger risk as FWI values larger than 45 are relatively common there.

Moreover, we studied the relationship between the March-September mean FWI and corresponding national forest fire statistics in Finland, Greece and Spain, see Table 2.

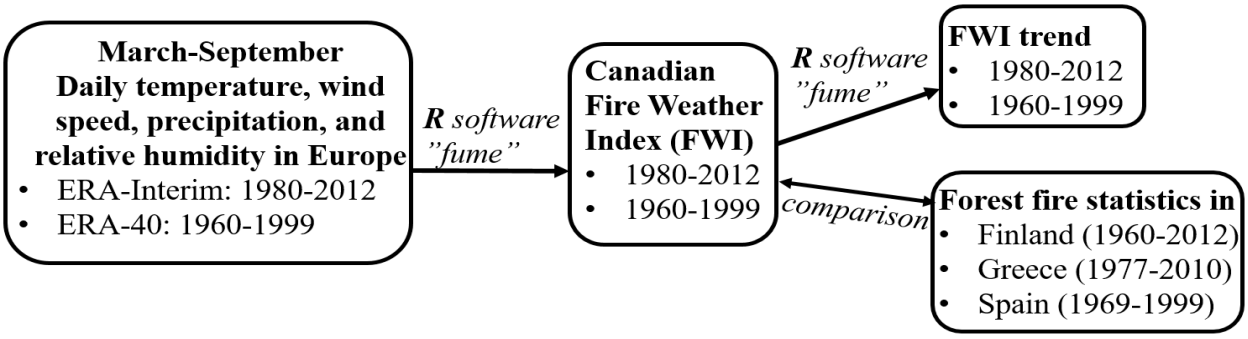

Figure 5 Flow chart showing the material and topics of Paper II. Figure by the author.

3.2.2 Projected changes in the geostrophic wind speeds (Paper I)

We explored the projected changes in the mean and extreme geostrophic wind speeds in Northern Europe from simulations of nine GCMs (Table 2). These models were selected from the Coupled Model Intergovernmental Project 3 (CMIP3) archive (Meehl et al. 2007) on the basis of spatial resolution being at least about 300 km (T42). We investigated three simulation periods: a baseline period for years 1971-2000, and future projections for 2046-65 and 2081-2100. The baseline period's (1970-2000) wind climate was compared to ERA-40 reanalysis (Uppala et al. 2005) (see Table 2) and the future scenarios were compared to the baseline period.

Table 2 Data used in Papers I and II.

Paper	Data	Source	Time period	Area
I	Daily mean • surface air pressure • surface temperature	ERA-40 reanalysis (Uppala et al., 2005)	1971–2000	Northern Europe
		GCMs (CMIP3, Meehl et al. 2007): BCCR-BCM2.0 CGCM3.1(T63) CNRM-CM3 ECHAM5/MPI-OM GFDL-CM2.1 IPSL-CM4 MIROC3.2(hires) MRI-CGCM2.3.2 NCAR-CCSM3	1971–2000 2046–2065 2081–2100	Northern Europe
II	• 12 UTC surface temperature • 24h precipitation sum Daily mean • Relative humidity • Wind speed	ERA-40 reanalysis (Uppala et al., 2005)	1961–1999	Europe
		ERA-Interim reanalysis (Dee et al., 2011)	1980–2012	Europé
	Forest fire statistics • Total burned area • Number of fires	Koutsias et al. (2013)	1977–2010	Greece
		EGIF, General statistics of Wildfires	1969–1999	Spain
		Finnish Forest research Institute (2010)	1960–2012	Finland

The future projections by the GCMs were forced by the Special Report on Emissions Scenarios (SRES, Nakicénovic et al. 2000), which are scenarios developed to estimate the unknown future greenhouse gas emissions and hence the atmospheric greenhouse gas concentrations of the 21st century. The SRES scenarios were used in the Third Assessment Report (IPCC, 2001) and in the Fourth Assessment Report (IPCC, 2007) of the Intergovernmental Panel on Climate Change (IPCC). The SRES-scenarios represent to two classes: consumer society scenarios with a more economic focus (A-scenarios) and scenarios aiming to sustainable development with a more environmental focus (B-scenarios). Here, three greenhouse gas scenarios B1, A1B, and A2 were utilized. The A2-scenario represents a rather pessimistic future with continuation of using fossil fuels resulting in high emissions and atmospheric CO_2 concentration of about 850 ppm by 2100 (see Table 3). The B1-scenario is a rather optimistic scenario with an emphasis on sustainable development with rapid development and introduction of environment friendly technology leading to decrease in greenhouse gas emissions with atmospheric CO_2 concentration of 550 ppm by 2100 (see Table 3). The A1B scenario represents an intermediate of the A2 and B1 scenarios with atmospheric CO_2 concentration of 720 ppm by 2100 (see Table 3).

In the IPCC Fifth Assessment Report (IPCC, 2013) updated greenhouse gas emission scenarios. Representative Concentration Pathway (RCP, van Vuuren et al. 2011) scenarios, were used. The RCPs are named after each scenario's corresponding change in the radiative forcing by year 2100 in comparison to pre-industrial radiative forcing. Thus, the additional radiative forcing by 2100, is in RCP2.6: 2.6 Wm^{-2}, in RCP4.5: 4.5 Wm^{-2}, in RCP6: 6.0 Wm^{-2}, and in RCP8.5: 8.5 Wm^{-2}. In RCP2.6 the global annual greenhouse gas emission peak between 2010–2020 and decline substantially thereafter leading to atmospheric CO_2 concentration of 421 ppm by 2100. In RCP 4.5

the global annual greenhouse gas emissions peak around 2040 and decline thereafter, resulting in atmospheric CO_2 concentration of 538 ppm by 2100. In RCP6, the emissions peak around 2080 and then decline leading to atmospheric CO_2 concentration of 670 ppm by 2100. In RCP8.5 the global greenhouse gas emissions continue to rise throughout the 21st century leading to high atmospheric CO_2, about 936 ppm (see Table 3). Projections based on the RCP scenarios were not available at the time **Paper I** was written and hence, in the discussion, we briefly compare our results to Ruosteenoja et al. (2019) who studied geostrophic wind speed changes over Europe with 21 GCMs from CMIP5 dataset forced by RCP4.5 and RCP8.5 scenarios, and to the review paper on storminess over the North Atlantic and northwestern Europe by Feser et al. (2015).

Comparison of SRES and RCP according to the atmospheric CO_2 concentration rates in 2100:

- A2 scenario is between RCP6.0 and RCP8.5,
- B1 is nearest to RCP4.5,
- A1B is nearest to RCP6.0.

Table 3 Projected atmospheric CO_2 concentrations in 2100 in selected SRES and RCP scenarios.

Scenario	**Projected Atmospheric CO_2 concentration in 2100**
A2	850 ppm
B1	550 ppm
A1B	720 ppm
RCP2.6	421 ppm
RCP4.5	538 ppm
RCP6.0	670 ppm
RCP8.5	936 ppm

In **Paper I** we investigated the geostrophic wind speeds rather than the simulated true surface wind speeds, as the geostrophic wind speeds are less affected by the model parametrization than the true surface wind speeds (Feser et al. 2015). The analysis focused on the high wind season from September to April. The annual maximum geostrophic wind speeds were investigated by comparing the 10-, 50-, and 100-year return periods of the baseline and future periods during the high wind season (September-April), see Figure 6. For the return period estimates we employed a free R software "extRemes" (Katz et al. 2005) using the Generalized Extreme Values - methodology (GEV, Coles, 2001) and the block maxima approach.

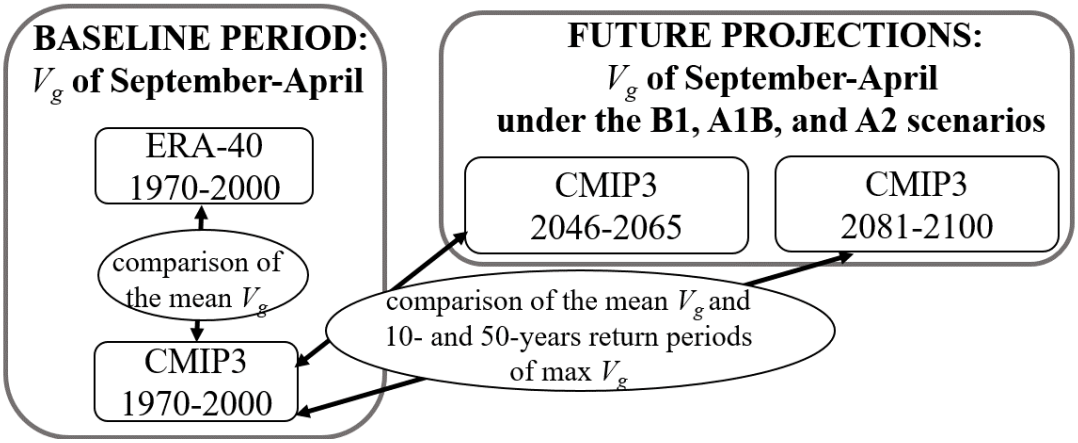

Figure 6 Flow chart showing the material and topics of Paper I. Figure by the author.

3.3 Statistical post-processing of ERFs (Paper V)

In **Paper V** we explored which stratospheric precursors favored statistically significantly weaker AO during the following 1-6 weeks in winters 1981-2016. Furthermore, we used the found precursors in post-processing winter surface temperature reforecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Vitart 2014) for forecast weeks 1 to 6 over Northern Europe.

3.3.1 Definition of the stratospheric wind indicator (SWI)

In **Paper V** (in review) the observed minimum AO indexes in the 1–6 weeks following the first Monday (as in 2017) in November-February in 1981-2016 were investigated. Inspired by the interactive comments by anonymous referees, we here show the observed mean (instead of minimum) AO indexes after all Mondays (as in 2017) in November-February in 1981-2016. More precisely, we examined the observed mean daily AO indexes (downloaded from the National Oceanic and Atmospheric Administration Climate Prediction Center) during the 1–2 weeks, 3–4 weeks, and 5–6 weeks following different phases of QBO (provided by the Free University of Berlin) and strengths of the daily ZMZW at 60°N and 10 hPa (provided by the National Aeronautics and Space Administration), see Table 4. As Scaife et al. (2014) demonstrated indicators of a more negative AO in the easterly QBO at level 30 hPa than in the westerly QBO phase at this level, we explored the AO index 1–6 weeks after following predictors:

- westerly QBO at 30 hPa, the *WQBO*,
- easterly QBO at 30 hPa, the *EQBO,*
- *EQBO* with the maximum of the monthly mean zonal wind components of the QBO between 70 hPa and 10hPa restricted to $7ms^{-1}$, $10ms^{-1}$, and $13ms^{-1}$,
- the daily ZMZW at 60° N and 10 hPa during the last 10 days of the previous month falling below its overall wintertime (November–March 1981–2016) 10^{th} percentile, corresponding a value of 3.8m/s, indicating a weak polar vortex already at the start of the forecast.

The statistical significance of the difference between the mean AO index following two different stratospheric situations, e.g., the *EQBO* and the *WQBO*, was determined using a two-sided Student's t-test with the null hypothesis that there is no difference. We used the most statistically significant predictors for weaker AO indexes observed 1–2 weeks, 3–4 weeks, and 5–6 weeks after these stratospheric situations, to define a stratospheric wind indicator (*SWI*) to be SWI_{neg}; otherwise, it was defined as SWI_{plain} for the beginning of each winter month (November–February) in 1981–2016 (see Decision tree in Figure 7).

Table 4 Data used in Paper V.

Paper	Data	Source	Time period	Area
V	• weekly mean surface temperature	ERA-Interim reanalysis (Dee et al., 2011)	1981–2016	Northern Europe
		ECMWF IFS Cycle 43r1 (Vitart, 2014) reforecasts	1997–2016	Northern Europe
	• daily surface AO index	NOAA, CPC	1981–2016	
	• daily ZMZW at 10 hPa	MERRA-2 (Rienecker et al. 2011)	1981–2016	60°N
	• monthly mean zonal wind at 70 hPa, 50 hPa, 40 hPa, 30 hPa, 20 hPa, 15 hPa and 10 hPa	Singapore radio soundings, Free University of Berlin (Naujokat 1986)	1981–2016	Singapore

3.3.2 Utilizing the SWI in post-processing ERFs

We investigated the observed and reforecasted surface temperature anomalies 1–2 weeks, 3–4 weeks, and 5–6 weeks after SWI_{neg} and SWI_{plain}. First, we calculated the observed two-week mean temperature anomalies of the ERA-Interim reanalyses (Dee et al. 2011) in January, February, March, November, and December in 1981–2016 in Northern Europe. Subsequently, we divided the observed two-week mean temperature anomalies to sets of anomalies, representing SWI_{neg} and SWI_{plain} according to stratospheric wind observations at the start of the forecast. Thereafter, we determined the statistical significance of the difference between the mean surface temperature anomalies after SWI_{neg} (TA_{SWIneg}) and SWI_{plain} ($TA_{SWIplain}$) using a two-sided Student's t-test with the null hypothesis that there is no difference between TA_{SWIneg} and $TA_{SWIplain}$.

We post-processed the ERFs of the ECMWF's Integrated Forecasting System (IFS Cycle 43r1, Vitart, 2014) which are run twice a week, on Mondays and Thursdays, in a horizontal resolution of 0.4 degrees. We first investigated weekly mean temperatures over Northern Europe in forecast weeks 1 to 6 of the forecasts initialized on Mondays. These raw reforecasts were mean bias-corrected (as in Buizza and Leutbecher 2015, eq. 7a, see Figure 7) by removing the mean bias computed from the ensemble reforecasts for the 20 years (1997–2016) depending on the forecast week date. For the 1997–2016 reforecasts, the mean bias was calculated considering 19 × 11 × 5 = 1045 ensemble reforecast members: 11 members' reforecast with initial dates defined by five weeks centered on the forecast week date for the 19 years reforecasts (1997–2016 excluding the reforecast year).

Those reforecasts, that were run in November-February 2017 for 1997–2016 (20 years × 17 runs = 340 reforecasts), were further post-processed by the observed *SWI* (see Figure 7). For the post-processing we first defined the *SWI* either SWI_{neg} or SWI_{plain} at the start of the forecast according to previous months' stratospheric wind conditions. According to the *SWI,* we added either TA_{SWIneg} or $TA_{SWIplain}$ to the ERA-Interim mean temperature during 1981–2016, corresponding to forecast weeks 1–2, 3–4, and 5–6 to get a SWI_{neg} and SWI_{plain} based mean temperatures, T_{SWIneg} and $T_{SWIplain}$, for weeks 1–2, 3–4, and 5–6, respectively. The T_{SWIneg} and $T_{SWIplain}$ were used in post-processing the ECMWF reforecasts' mean bias-corrected ensemble members, T_{BC}, by calculating T_{SWI_BC} for SWI_{neg} as follows:

$$T_{SWI\_BC} = (1 - k_{SWI}) * T_{BC} + k_{SWI} * T_{SWIneg} \quad (1)$$

And for SWI_{plain},

$$T_{SWI\_BC} = (1 - k_{SWI}) * T_{BC} + k_{SWI} * T_{SWIplain} \quad (2)$$

where T_{SWI_BC} was a post-processed ensemble member. k_{SWI} was the weight of the T_{SWIneg} or $T_{SWIplain}$, which was tested between 0–1 and defined according to the best improvement in the skill scores of the post-processed forecast. By Eq. (1) and Eq. (2), we adjusted each ensemble member with the same weight, and hence, the original spread of the ECMWF reforecasts remained unchanged.

The two-week averages of both the mean bias-corrected and the post-processed surface temperature forecasts over Northern Europe were verified against ERA-Interim 1997–2016 temperature re-analyses (Dee et al. 2011) by a commonly used measure for probabilistic forecasts:

the continuous ranked probability score (CRPS, Hersbach 2000). The skill scores of the CRPS, the CRPSS, were calculated by comparing the reforecasts to the climatological forecast of years 1981–2010 of the ERA-Interim data.

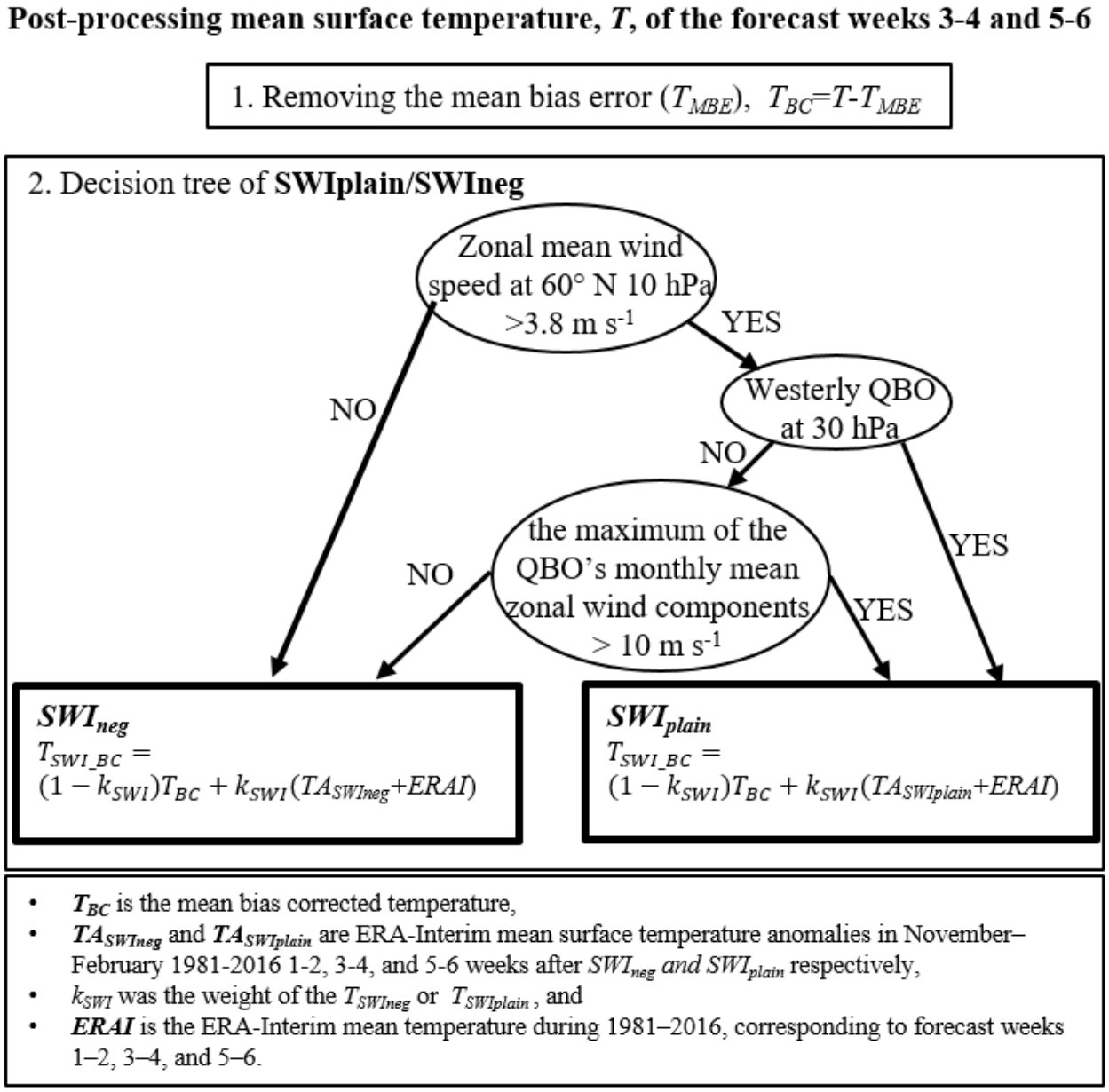

Figure 7 Flow chart showing the post-processing of the surface temperature forecasts. Figure by the author.

4. RESULTS

4.1 Glacial climate in Europe

4.1.1 Downscaled temperature and precipitation (Paper III)

The GAM fitted to downscale annual mean precipitation in Europe (in **Paper III** GAM_Western Eurasia) in both LGM and present day climate, showed good correlation (0.84) with the fitted data and was able to explain 85% of the spatial variance of the total precipitation. The GAM fitted to downscale annual mean temperature in Europe (in **Paper III** GAM_Western Eurasia) in both LGM and present day climate, also showed good correlation (0.99) with the fitted data and was able to explain 98% of the spatial variance of the annual mean temperature.

The GAMs for downscaling temperature were used for downscaling the annual mean temperature between 11 and 0 kyr BP using as input predictors the 11 to 0 kyr BP data from the CLIMBER-2-SICOPOLIS last glacial cycle simulation by Ganopolski et al. (2010). The predictions are plotted in Figure 8, together with pollen-based reconstructions of annual mean temperature and the bilinear interpolation of the annual mean surface temperature simulated by CLIMBER-2 at two locations in Northern Europe during the Holocene, between 11 and 0 kyr BP. The shapes of the CLIMBER-2 simulated and GAM downscaled temperature curves were generally consistent with temperature estimates based on paleoclimatic reconstructions. As for the absolute values, the GAM-downscaled CLIMBER-2 temperature showed better agreement with the pollen-based reconstructions than the bilinear interpolation of the CLIMBER-2 annual mean surface temperature.

The GAMs which were calibrated only by the recent past and the LGM data (GAM_Western Eurasia) were used for predicting annual mean temperature and precipitation for 44 kyr BP using as input predictors the 44 kyr BP data from the CLIMBER-2-SICOPOLIS last glacial cycle simulation by Ganopolski et al. (2010). The predictions are plotted in Figure 9, together with output of CLIMBER-2 and RCA3 regional climate model simulation output by Kjellström et al. (2010) representing 44 kyr BP annual mean temperature and precipitation. Over Eastern Europe the GAM (Figure 9c) predicted similar temperature spatial patterns as the RCA3 (Figure 9b), whereas over Western Europe, there were more differences in details. Over Finland the differences between the RCA3 simulated surface temperature and the CLIMBER-2 surface temperature downscaled by GAM mainly result from the difference of ice sheet extent in SICOPOLIS and RCA3: the Fennoscandian ice sheet in SICOPOLIS reached Central Finland during 44 kyr BP, whereas in the RCA3 simulation the ice sheet reached only some parts of Northern Finland (Figure 1 in **Paper III**). The CLIMBER-2 annual precipitation downscaled by our GAM in Figure 9f were similar to the RCA3 simulation output in Figure 9e in many parts of Eastern and Central Europe, however, not predicting as high precipitation amounts in mountain regions and as low in the Mediterranean Sea.

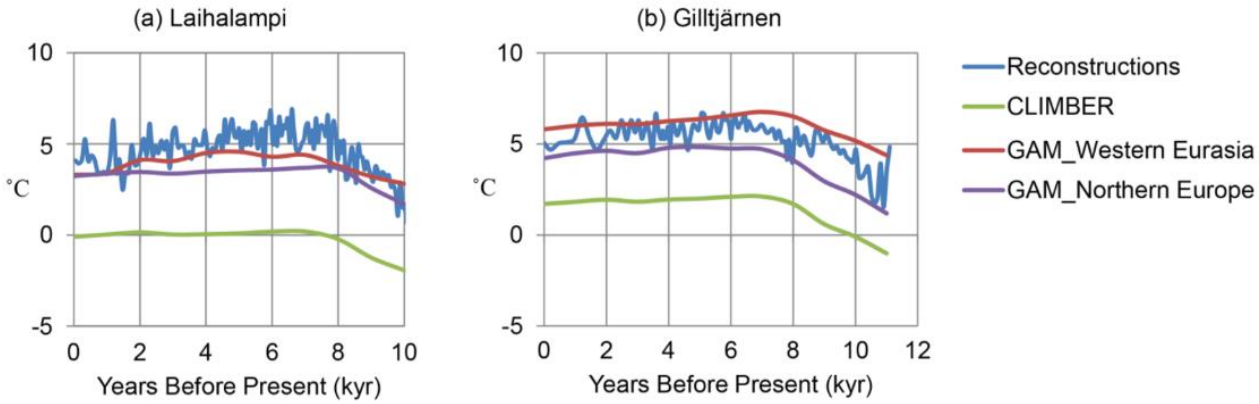

Figure 8 Comparison of simulation data and polled-based reconstructions of annual mean temperatures from (a) Laihalampi in Finland and (b) Gilltjärnen in Sweden. The blue curves represent reconstructions by Heikkilä and Seppä (2003) for Laihalampi, and by Antonsson et al. (2006) for Gilltjärnen. The green contours are interpolations from CLIMBER-2 simulations by Ganopolski et al. (2010), the red contours represent annual mean temperatures downscaled by the GAM for Europe. Figure from ***Paper III****, © Authors 2014. CC Attribution 3.0 License.*

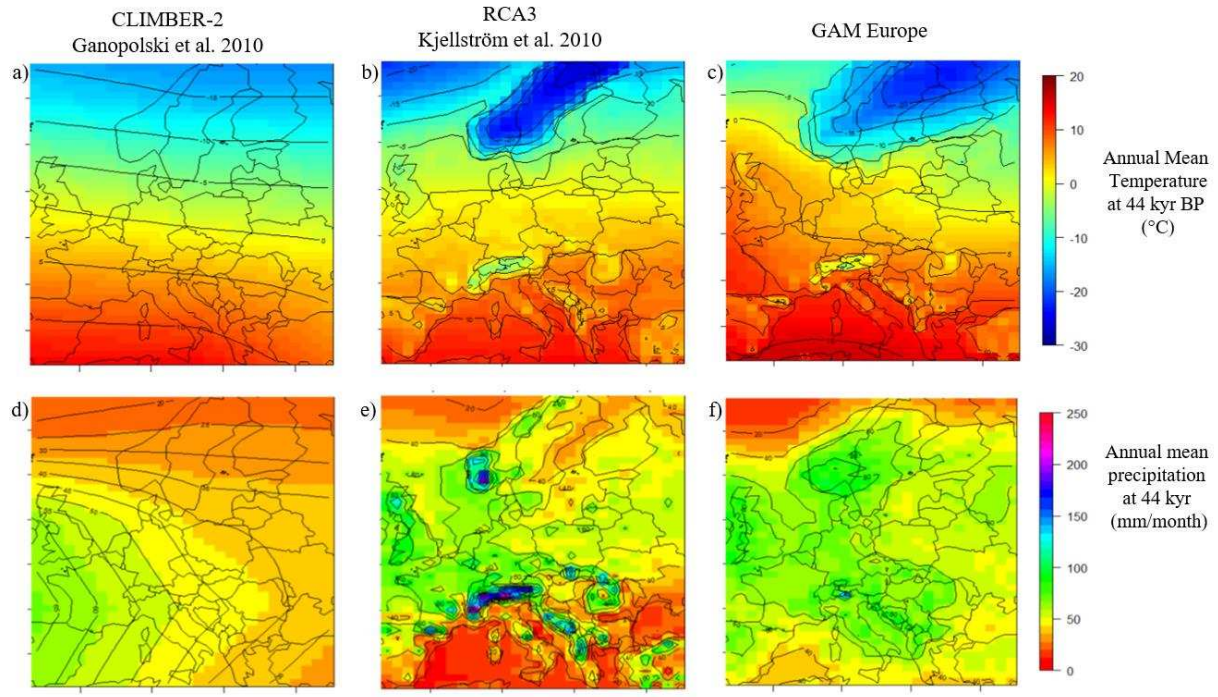

Figure 9 Annual mean temperature at 44 kyr BP, a) as simulated by the global model CLIMBER-2 (Ganopolski et al. 2010), b) as simulated by the regional model RCA3 (Kjellström et al. 2010), and c) as predicted by the GAM for Europe. Annual mean total precipitation at 44 kyr BP, d) as simulated by the global model CLIMBER-2 (Ganopolski et al. 2010), e) as simulated by the regional model RCA3 (Kjellström et al. 2010), and f) as predicted by the GAM for Europe. The data of CLIMBER-2 and RCA3 have been bi-linearly interpolated on to a 1.5º × 0.75º resolution. Figure modified from figures S2.1 and S2.2 in the supplementary material of ***Paper III****, © Authors 2014. CC Attribution 3.0 License.*

4.1.2 Human population size and range simulations (Paper IV)

The GAMs calibrated in **Paper III** were used for downscaling annual mean temperature and precipitation between 30 and 13 kyr BP in Europe using as input predictors the 30 to 13 kyr BP data from the CLIMBER-2-SICOPOLIS last glacial cycle simulation by Ganopolski et al. (2010). The downscaled temperature and precipitation data were used to estimate the potential evapotranspiration (Figure 10D), mean temperature of the coldest month (Figure 10E), and water balance (Figure 10F) in Europe between 30 and 13 kyr BP. These climatic parameters were used to estimate hunter-gatherer population size and density in Europe between 30 and 13 kyr BP by the population calibration model developed in **Paper IV**. Figure 10 depicts simulated hunter-gatherer population size and density, the archaeological population proxy, and the GAM based climate estimates between 30 and 13 kyr ago in Europe.

According to the simulations, the potentially inhabited land area in Europe between 30 and 13 kyr BP, extended about 500 km south of the Fennoscandian ice sheet margin from the Iberian Peninsula to southern parts of modern Ukraina and European Russia with a southwest-northeast gradient of decreasing population densities. During the severe cooling and expansion of the ice sheets between 30 and 23 kyr ago, when the simulated mean temperature of the coldest month in Europe decreased about 8 degrees in 7,000 years, the simulated human population size declined from 330,000 people to 130,000 people (Figure 10A). During the very end of the last glacial between 19 and 13 kyr ago, when the simulated mean temperature of the coldest month in Europe rose about 12 degrees, the simulated population size grew from 150,000 people to over 400,000 people in Europe (Figure 10A). The simulated human population size changes were compared to estimates of the human population size changes based on archaeological population proxy data (Figure 10C) and significant correspondence was found, suggesting that climate indeed was a major driver of population dynamics between 30 and 13 kyr ago.

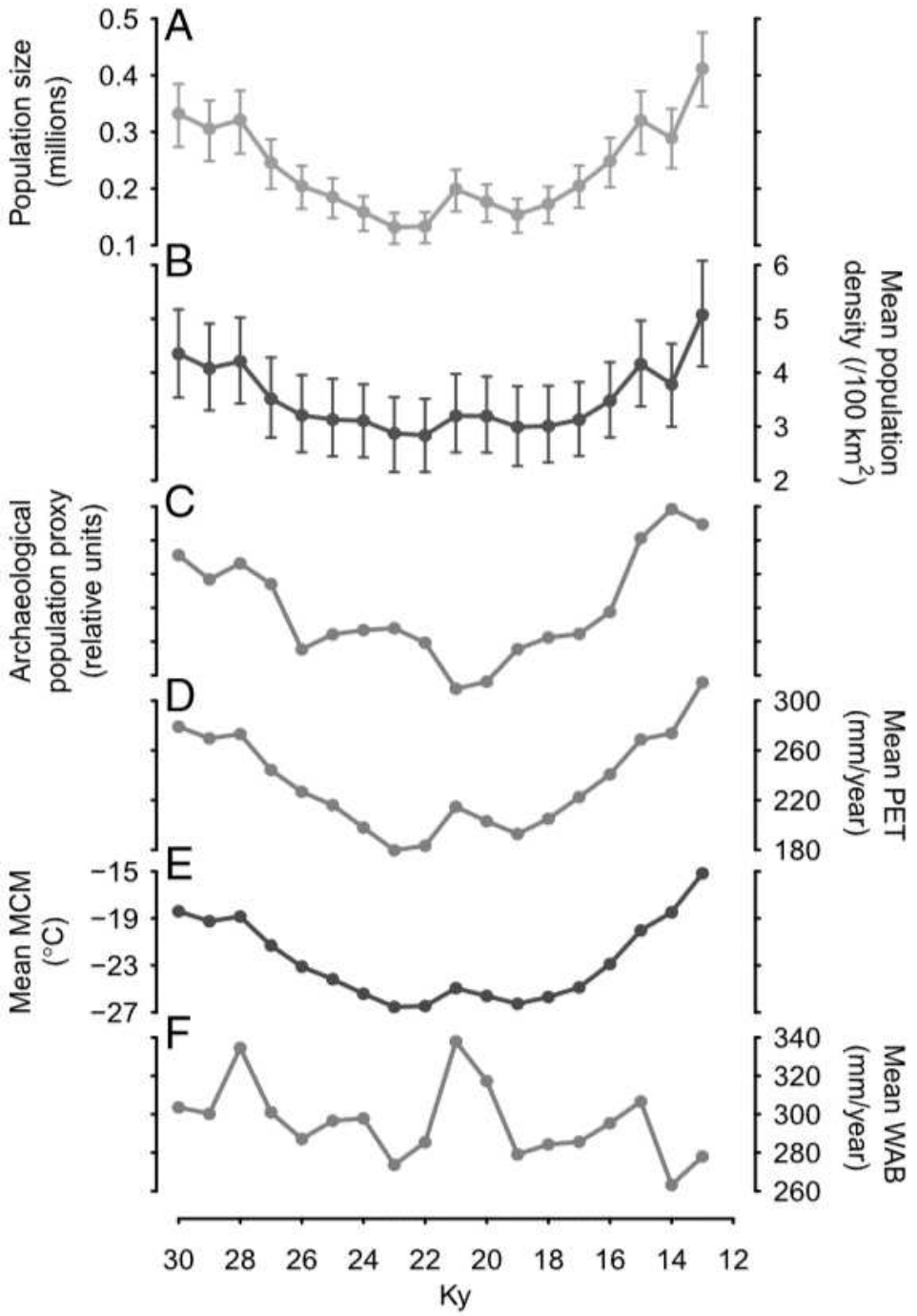

Figure 10. Comparisons between simulated hunter-gatherer population size and density, the archaeological population proxy, and paleoclimatic simulations between 30 and 13 kyr ago in Europe. (A) Simulated human population size in Europe. Error bars show the resampling-based confidence limits (95%). (B) Simulated mean density in the inhabited area of Europe. Error bars show the resampling based confidence limits (95%). (C) Archaeological population size proxy based on the taphonomically corrected number of dates. (D) European mean of simulated potential evapotranspiration. (E) European mean of simulated mean temperature of the coldest month. (F) European mean of simulated water balance. D-F are based on the downscaling of the CLIMBER-2 climate model. Figure from ***Paper IV****.*

4.2. Human-induced climate change impacts in Europe

4.2.1 Response of the FWI (Paper II)

The March-September mean values of FWI variated relative largely from year-to-year between years 1960 and 2012 in whole Europe and in the specified sub-areas. In Southern Europe the FWI had relatively high values in early 1960s (around 35), lower values in the early 1970s (between 25 and 30), and rising trends thereafter (between 35 and 40 in years 2000-2012). In Eastern Europe the FWI had relatively high values in early 1960s (between 10 and 15), lower values in the 1970s (between 5 and 10), and rising trends thereafter (mostly between 8 and 13 in years 2000-2012).

Figure 11 shows the trends of mean FWI in each grid point over Europe using the ERA-40 for 1960-1999 and ERA-Interim for 1980-2012. ERA-40 based mean March-September FWI showed some statistically significant upward trend for parts of central Europe and Balkan Peninsula. In the area north of the Black Sea, slightly decreasing trends were found in ERA-40 based mean FWI (1960-1999), whereas the later period, 1980-2012 of ERA-Interim, showed there increasing trends. ERA-Interim based FWI showed statistically significant rising trends also in large areas of Eastern Europe, France, and Iberian Peninsula.

For Northern Europe the trend analyses for both ERA-Interim and ERA-40 based mean March-September FWI showed no significant trend. For southern and eastern Europe in 1980-2012, the trend analyses for ERA-Interim based mean March-September FWI showed significant upward trend at the 99% level. For southern (western and eastern) Europe the trend was significantly upward at the 99% (90%) level also when both ERA-Interim and ERA-40 data were used, i.e., in years 1960-2012. In contrast, the ERA-40 based mean March-September FWI did not show any significant trend in any of the specified areas for the period of 1960-1999.

The time series for the number of days when FWI exceeds threshold 20 or 45 were similar to the time series of the mean FWI based on ERA-Interim data. ERA-Interim based number of days where FWI>20 showed upward trend for southern Europe significant at the 99% and for Eastern Europe significant at the 95% level. For the number of days where FWI>45 showed upward trend for southern Europe significant at the 99% level. For Western and Northern Europe there was no trend found.

Moreover, we compared the March-September mean FWI and corresponding national forest fire statistics in Finland, Greece and Spain. We found that in these study areas, the correlations between the FWI and the size of the area burned (correlation coefficients between 0.6 and 0.7) were higher than with the correlations between the FWI and the number of forest fires.

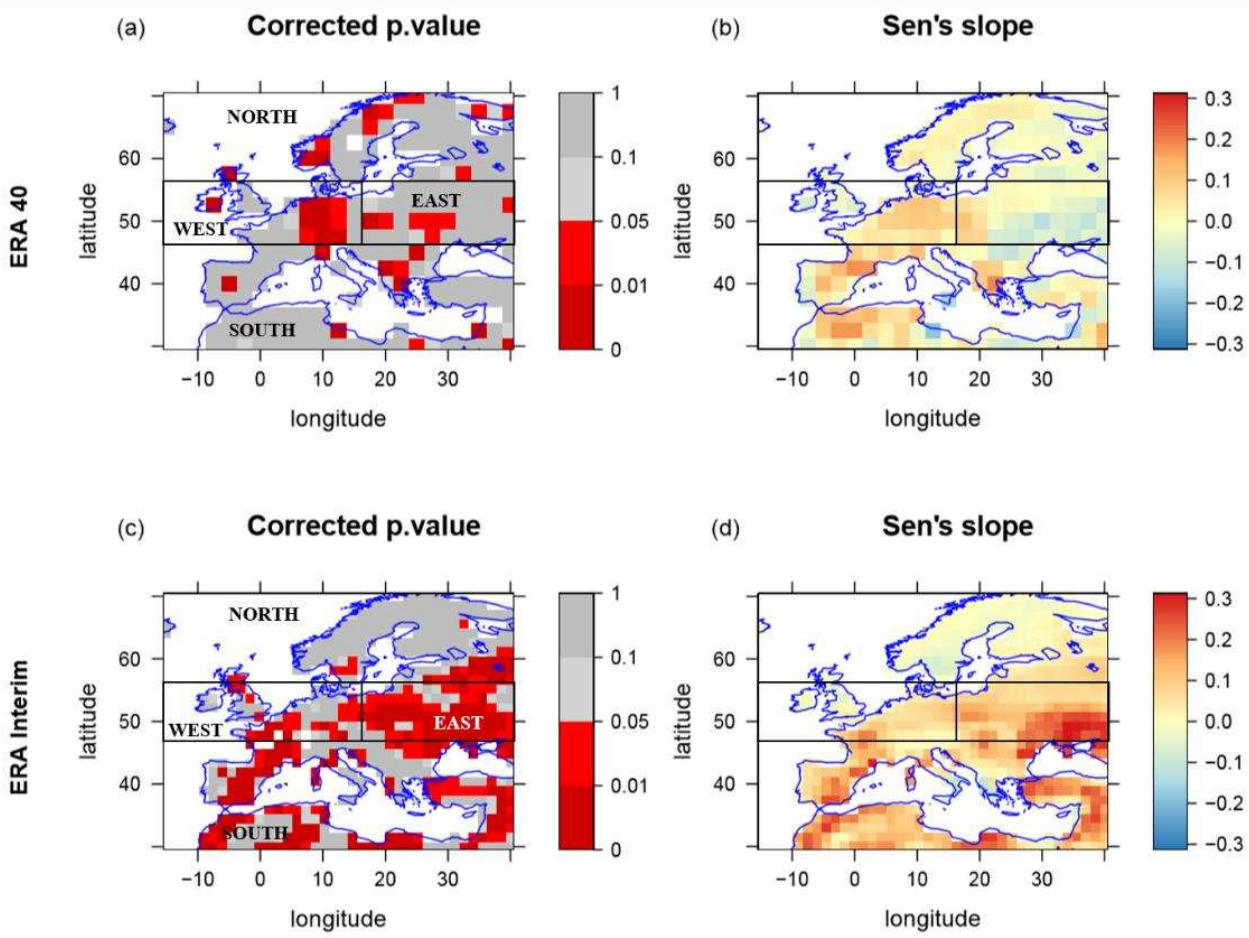

Figure 11 Trend of March-September mean FWI calculated using the ERA-40 data set for 1960-1999 (b) and the ERA-Interim data set for 1980-2012 (d). The statistical significance (α) is shown at levels 0.01, 0.04, and 0.1 for ERA-40 (a) and ERA-Interim (c). Figure modified from Figures 1 and 3 in ***Paper II****, © Authors 2014. CC Attribution 3.0 License.).*

Impact models such as FWI are used to describe the impact of climate change. Here the upward trend of the FWI in southern Europe can be explained by the observed warming and decrease in precipitation. For northern Europe there was no trend found in the mean FWI in 1960-2012. In Northern Europe the temperatures have also risen, but the precipitation has not decreased during this period. The temporal variations of FWI found in **Paper II** depict the long-term temporal variations of European climate. Also in the future - although climate is warming - the long-term climate variations will remain and periods with lower FWI can be experienced even in areas characterized by an increasing trend of FWI.

4.2.2 Projected changes of the geostrophic wind speeds (Paper I)

The 9-GCM simulated mean geostrophic winds of the baseline climate (1970-2000) were compared with the observation-based ERA-40 reanalysis. In this comparison, the bias of the modelled geostrophic winds was generally fairly small, around 0.5ms^{-1}, and the seasonal cycle of the modelled geostrophic winds was found to be reproduced reasonably well.

The multi model mean projections of the September-April mean geostrophic wind speeds in Northern Europe showed similar patterns of increase and decrease, but with slightly different values depending on the greenhouse gas forcing (Fig. 3 in **Paper I**). In all greenhouse gas emission scenarios the projections showed statistically significant 2-4% decrease in the mean geostrophic wind speeds over the Norwegian Sea already by the period 2045-2065. In addition, by 2045-2065, the projections forced by the A1B and A2 scenarios showed statistically significant 1-2% increase in the mean geostrophic winds in parts of southern Baltic Sea. Further, by 2100 the projections forced by A1B and A2 scenarios showed statistically significant 2-4% increase in the mean geostrophic wind speeds in northwestern parts of Russia and southern parts of the Baltic Sea. The projections forced by the B1 scenario did not indicate any statistically significant increase in the mean geostrophic wind speeds anywhere in Northern Europe during the ongoing 21st century.

Figure 12 shows the projected changes in the September-April extreme geostrophic wind speeds for the ongoing century under different greenhouse gas emission scenarios. The patterns of increase and decrease of extreme geostrophic winds were similar to the changes in the mean geostrophic wind speeds and the changes were more pronounced with higher greenhouse gas concentrations. Statistically significant decrease in the extreme geostrophic winds were projected over the Norwegian Sea in the A1B and B1 greenhouse gas emission scenarios (2-6% decrease) until year 2065, and in the A1B and A2 scenarios (2-8% decrease) until the end of the ongoing century. Moreover, in the A1B and A2 emission scenarios a statistically significant 2-4% increase in the extreme geostrophic winds were projected in southern parts of the Baltic Sea, a few parts of Finland, and in Northwestern parts of Russia until the end of the ongoing century. In the B1 emission scenario there was again no statistically significant increase in extreme wind speeds projected during the ongoing century.

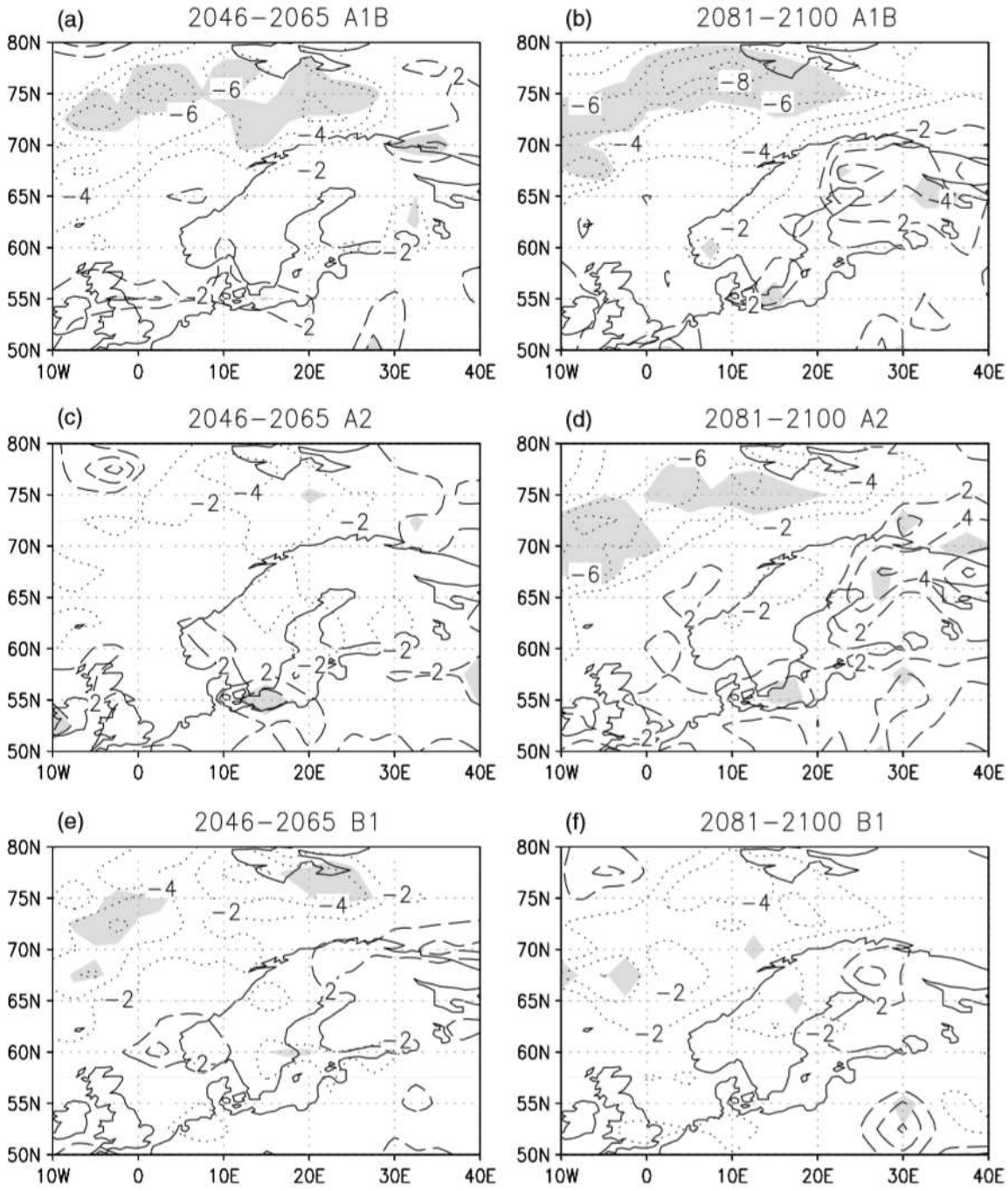

Figure 12 The changes in the 10-year return level estimates from 1971-2000 to 2046-2065 (on the left) and from 1971-2000 to 2081-2100 (on the right) for individual SRES- scenarios as an average of six GCM simulations. From the top the 6GCM mean (a) and (b) A1B; (c) and (d) A2; (e) and (f) B1. Shading indicates the areas where the change is statistically significant at the 95% level. Figure from ***Paper I****, © 2011 Royal Meteorological Society.*

4.3 Statistical post-processing of ERFs (Paper V)

4.3.1 Stratospheric winds as precursors of the surface

Figure 13 shows boxplots of the observed means of the daily AO indexes 1–2 weeks, 3–4 weeks, and 5–6 weeks after different phases of QBO and restrictions in the strength of the stratospheric winds in 1981–2016. The first box (brown) represents the mean AO indexes after all the cases in 1981–2016 November–February. The second and third boxes show the mean AO indexes after easterly (*EQBO*, blue) and westerly (*WQBO*, pink) QBO at the 30 hPa level, respectively. The p-value written below each boxplot pair indicates the likelihood of such a pair of distributions arising from a random sampling of a single distribution as given by a Student's t-test, i.e., p-values less than 0.05 indicate that the means of the data sets differ significantly at the 95% level of confidence. The median and the mean of the mean AO indexes 1–2 weeks, 3–4 weeks, and 5–6 weeks after *EQBO* were lower than after *WQBO*. The *EQBO* (blue) box shows all the cases of *EQBO* with no restriction in the QBO's monthly mean zonal wind components, whereas the fourth, the sixth, and the eighth blueish boxes show the mean AO indexes after *EQBO* with all the QBO's monthly mean zonal wind components between levels 70…10hPa being below 13 m/s, 10 m/s, and 7 m/s, respectively. Restricting the *EQBO* cases by a maximum of the QBO's monthly mean zonal wind components in levels 70…10hPa decreased the median of the mean AO during the following 1–2, 3–4, and 5–6 weeks.

The 10th box (yellow) in Figure 13 shows the mean AO indexes after cases the daily ZMZW at 60° N and 10 hPa was below its 10th percentile (3.8m/s) during the last 10 days preceding the start of the forecast, corresponding to cases with weak polar vortex already at the start of the forecast. The observed mean AO index was statistically significantly weaker at the 99% confidence level 1–2 weeks, 3–4 weeks, and 5–6 weeks after the daily ZMZW at 60° N and 10 hPa had been below its overall wintertime 10th percentile.

Aiming to select stratospheric precursors indicating weak AO with the greatest statistical significance, we defined the *SWI* to be negative in cases when the QBO was easterly at 30 hPa and the QBO's monthly mean zonal wind components in levels 70…10hPa were weaker than 10m/s or if the daily ZMZW at 60° N and 10 hPa during the last 10 days preceding the start of the forecast fell below its overall wintertime 10th percentile. In other cases, the *SWI* was defined as plain 2016 (see Decision tree in Figure 7).

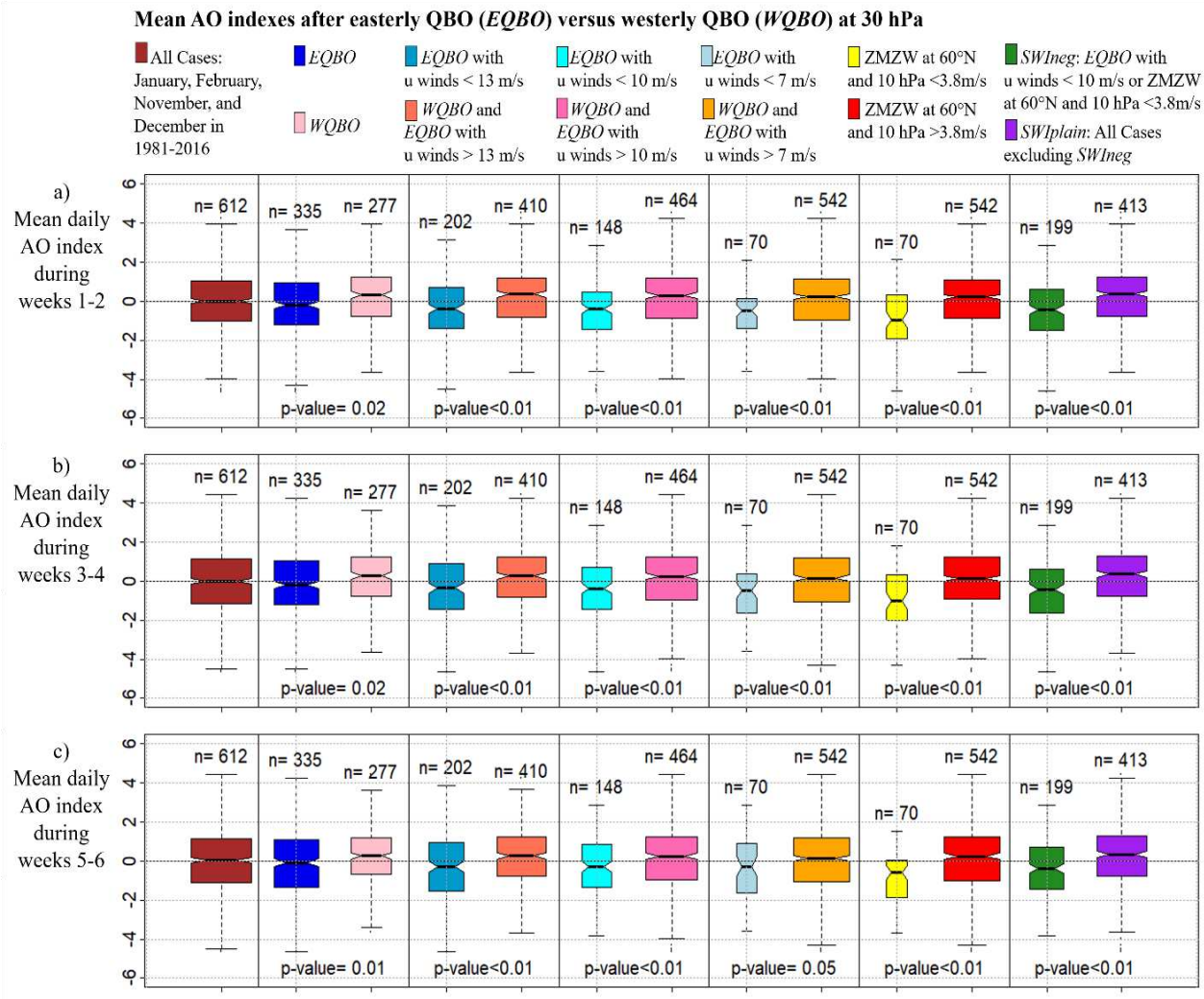

Figure 13 Observed mean daily AO index a) 1–2, b) 3–4 and c) 5–6 weeks after different stratospheric situations. The line dividing each box into two parts shows the median of the data, the ends of the box show the lower and upper quartiles, and the whiskers represent the highest and the lowest values excluding outliers. The n written above each box indicates the number of observations in each group. The widths of the boxes have been drawn proportional to the square-roots of n. The p-value written below each boxplot pair indicates the likelihood of such a pair of distributions arising from a random sampling of a single distribution as given by a Student's t-test, i.e., p-values less than 0.05 indicate that the means of the data sets differ significantly at the 95% level of confidence. The notches of each side of the boxes were calculated by R boxplot.stats. If the notches of two plots do not overlap, this is 'strong evidence' that the two medians differ (Chambers et al., 1983, p. 62). ZMZW=zonal mean zonal wind.

Figure 14 shows the observed (periods 1981–2016 and 1997–2016) and model forecasted (the period 1997–2016) mean temperature anomalies of the weeks' 1–2, 3–4, and 5–6 in November–February after SWI_{neg} and SWI_{plain}. The observations showed on average lower mean temperatures 1-2, 3–4 and 5–6 weeks after SWI_{neg} (Fig. Figure 14 a-c and Figure 14 g-i). This cold anomaly was forecasted well in the reforecasts for weeks 1–2 (Figure 14 m), and it was underestimated in reforecasts for weeks 3–4 (Figure 14 n) and 5–6 (Figure 14 o). Further, the observations showed on average higher mean temperatures 1–2, 3–4, and 5–6 weeks after SWI_{plain} (Figure 14 d-f and Figure 14 j-l). This warm anomaly was forecasted well in the forecasts weeks 1–2 (Figure 14 p), and it was underestimated in forecast weeks 3–4 (Figure 14q) and 5–6 (Figure 14 r).

The mean temperature anomalies 3-6 weeks after SWI_{neg} (Figure 14 b-c) and SWI_{plain} (Figure 14 e-f) during 1981–2016 were statistically significantly different using a Student's t-test, with anomalously cold surface temperatures more common 3-6 weeks after SWI_{neg}. When examining the years 1997–2016 (Figure 14 h-i and Figure 14 k-l), which was the reforecast period, the temperature anomalies were of the same sign than during the longer 1981–2016 period (Figure 14 b-c and Figure 14 e-f) but weaker and not statistically significant in weeks 5-6 all over the Northern Europe.

The stronger warm anomalies after SWI_{plain}, and weaker cold anomalies after SWI_{neg} in 1997-2016 (Figure 14 g-l) in comparison to 1981-2016 (Figure 14 a-f) are qualitatively in line with the observed annual mean temperature rise in Finland (e.g., Mikkonen et al. 2015)

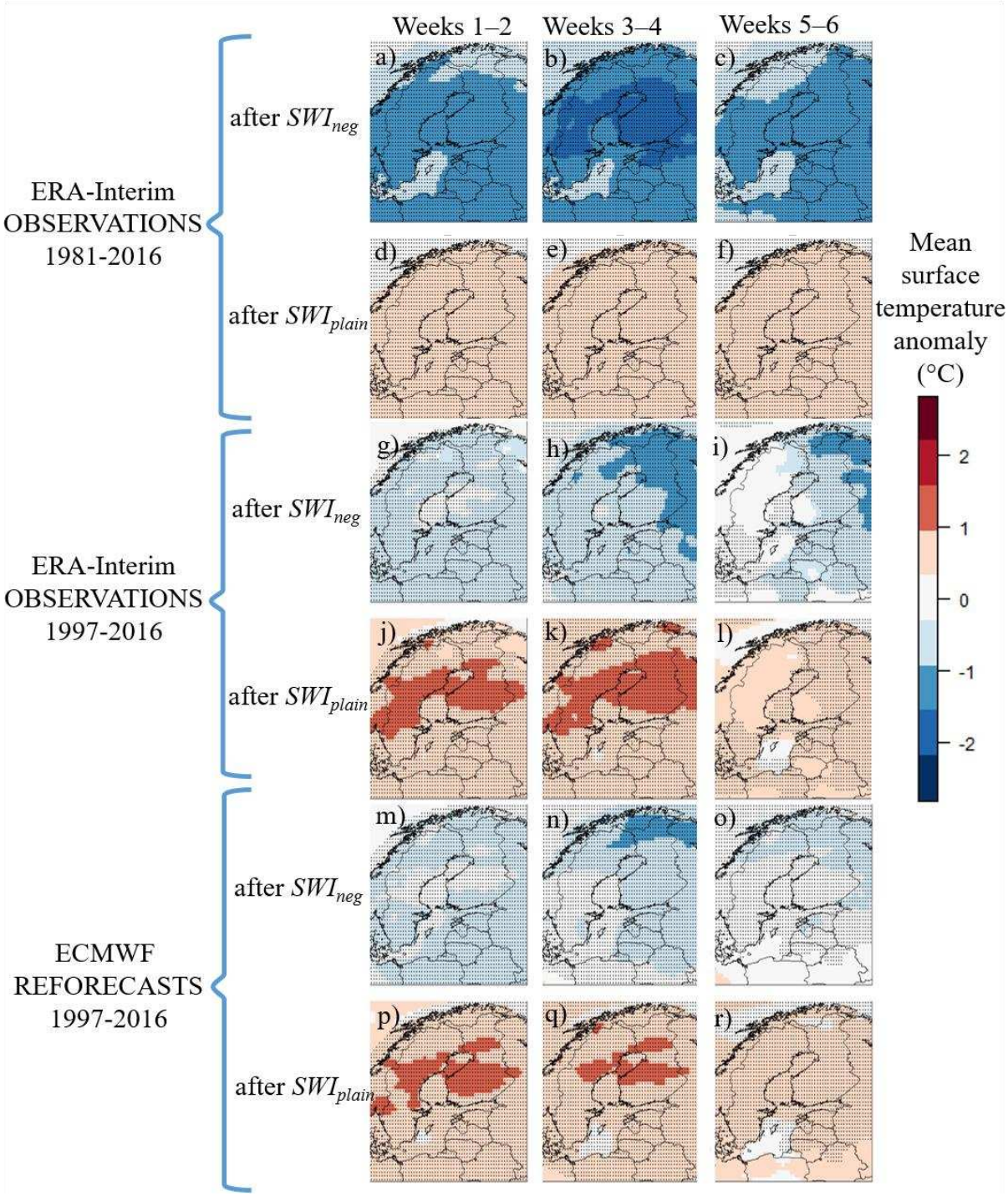

Figure 14 ERA-Interim observed (a-l) and ECMWF reforecasted (m-r) mean temperature anomalies in comparison to 1981-2016 mean during boreal winters (November-February) in cases the previous month's SWI was negative (SWI_{neg}, covering about 30% of the winter months) or plain (SWI_{plain}, covering about 70% of the winter months). The dotted areas represent the 95% level of confidence where the means of surface temperature anomalies after SWI_{neg} and SWI_{plain} differ significantly.

4.3.2 SWI in post-processing ERFs

The mean temperature anomalies in Figure 14 (a-f) for Northern Europe were used for the *SWI* based post-processing as by Eq. (1) and (2). The best median CRPSS was achieved by k_{SWI}=0.4, for forecast weeks 3–4 and by k_{SWI}=0.6 for forecast weeks 5–6. Figure 15 shows the forecasts skill of the mean temperature of the forecast weeks 3–4 and weeks 5–6 forecasted by the mean bias-corrected reforecasts alone (Figure 15 a-b) and by the mean bias-corrected reforecasts together with the *SWI* based post-processed forecast (Figure 15 c-d). By using the *SWI* based post-processing to the ECMWF forecasts, the CRPSSs for weeks 3–4 and weeks 5–6 were slightly improved and the area of these forecasts being significantly better than just the climatological forecast was expanded.

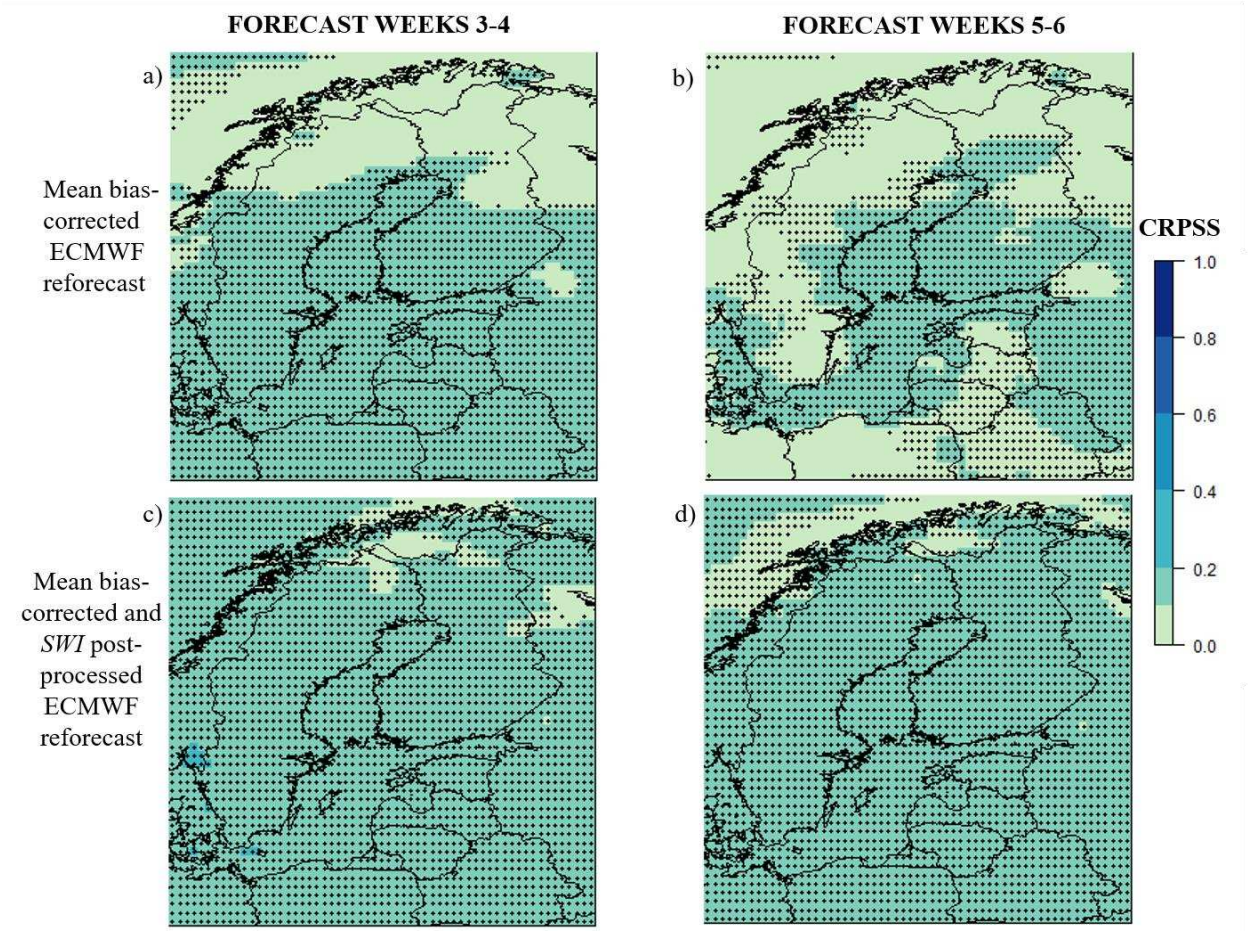

Figure 15 Expected CRPSS of forecast weeks 3–4 and 5–6 of the ECMWF's mean temperature reforecasts for November–February 1997–2016 after mean bias-correction (a-b) and after both mean bias-correction and the SWI based post-processing (c-d). ERA-Interim climatology of 1981–2010 was used as the reference. The dotted areas represent the 95% level of confidence that the CRPSS is above zero.

5. DISCUSSION

Spatial patterns of the annual mean surface temperature and precipitation in Europe during the last glacial cycle were in **Paper III** reproduced with GAMs by downscaling the coarse resolution EMIC simulations by Ganopolski et al. (2010). These GAM downscaled annual mean surface temperature and precipitation were compared to pollen-based proxy reconstructions representing 11 to 0 kyr BP (Heikkilä and Seppä (2003) and Antonsson et al. (2006)) and a GCM based time slice simulation of glacial climate representing 44 kyr BP (Kjellström et al. 2010). Although some of the detailed spatial features were not fully captured in regions with large spatial variation of topography, the main spatial patterns were relatively well captured. In **Paper III** the GAMs for downscaling were calibrated by both present day and GCM simulated LGM climate data, which widened the calibration range of the used GAMs in comparison to Vrac et al. (2007) who downscaled LGM climate by GAMs calibrated by present day climate data only.

The climate seems to have been a major driver of the hunter-gatherer population size in Europe during the end of the last glaciation, as the climate based modelled human population size changes in Europe between 30 and 13 kyr BP, in **Paper IV**, correlated well with corresponding independent archaeological estimates of human population size changes (Vermeersch, 2005). Further environmental factors affecting the hunter-gatherer population density have been recently studied by Tallavaara et al. (2018), who studied the effect of the climate based net primary production, the biodiversity richness and the environmental pathological stress on hunter-gatherer population density patterns. Of these three environmental factors, in the high and mid-latitudes the higher net primary production and richness of the biodiversity had the greatest (positive) influence on the hunter-gatherer population densities, whereas in the tropics and sub-tropics the environmental pathogen stress had the strongest limiting effect on the hunter-gatherer abundance.

Under the on-going human-induced climate change, the trend of climate-based forest fire risk (FWI), studied in **Paper II,** has risen in Eastern and Southern Europe during 1960-2012. It is important to note, however, that in addition to climate, important factors contributing to forest fire risk are: human activities and the effectiveness of fire detection and suppression systems (Bowman et al. 2009). The climate-based forest fire risk has been estimated to increase also in the Northern Europe by the end of the ongoing century (Lehtonen et al. 2014, Mäkelä et al. 2014, Lehtonen et al. 2016).

According to projections by 9 GCMs (of CMIP3) under SRES A2, B1 and A1B scenarios, in **Paper I,** the changes in the mean and extreme winds in Northern Europe are going to be small during the ongoing century. This is in line with later studies, e.g., with Ruosteenoja et al. (2019) who used 21 GCMs (of CMIP5) under moderate-emission RCP4.5 and large-emission RCP8.5 scenarios. Ruosteenoja et al. (2019) found that although the mean and extreme winds are not projected to increase much, the frequency of strong westerly winds are projected to rise and the frequency of strong easterly winds are projected to decrease. Feser et al. (2015) analyzed a number of pre-CMIP3, CMIP3 and CMIP5 model studies to assess potential future changes in storm climate over the North Atlantic and Northwestern Europe; they found no clear trend in storm numbers, but most of the analyzed studies agreed on an increase in storm intensity for the future. However, whether or not the mean and extreme winds are going to increase, the damages might

anyhow increase, for example on forests: as the winters are expected to be milder and the ground less frozen, storms might cause bigger timber losses in winter (Gregow et al., 2011b).

In **Paper V** novel information about the statistical relationship of the QBO and the AO index was found. In addition to the previously demonstrated more negative AO during easterly QBO in comparison to westerly QBO at 30 hPa (Scaife et al. 2014), we found that the maximum strength of the QBO's monthly mean zonal wind components during the easterly QBO affected the observed mean AO index 1-6 weeks later. This connection was used in post-processing 3-6 weeks surface temperature forecasts in Northern Europe and the skill scores of these forecasts were thereby slightly improved. As Domeisen (2019a, 2019b) showed that all of the current S2S database's ERF prediction systems have an enhanced prediction skill in the extratropical stratosphere in comparison to the troposphere, the next step would be looking for the stratospheric signals directly from the forecast model and use them in post-processing surface temperature forecasts. Considering QBO and the ongoing global warming: according to model simulations by Kawatani et al. (2011) in the atmosphere with doubled CO_2 concentration, the amplitude of the QBO weakened and the length of the QBO's phase increased. However, based on model simulations by Schirber et al, (2015) the amplitude and length of the QBO's phases were sensitive to model parametrizations, in particular to parametrizations of gravitational waves. Also the stratospheric polar vortex might change in the warming climate. Manzini et al. (2014) reported about a possibly weakening winter stratospheric polar vortex in the warming climate (CMIP5 simulations). However, Ayarzagüena, et al. (2018), reported that this might not have any significant effect on the frequency of the SSWs in the future climate. All in all, the future changes in both the stratospheric polar vortex and the QBO are somewhat uncertain.

CONCLUSIONS

In this thesis I have demonstrated use of versatile datasets and methods for assessing climate variation and change. As well, a few climate change impacts have been studied.

The main conclusions of this work are as follows:

1. In **Paper III**, the GAMs showed to be suitable in downscaling annual mean temperature and precipitation of the CLIMBER-2 EMIC's coarse resolution simulation of the last glacial cycle (Ganopolski et al. 2010). GAMs that were fitted by both LGM (Brady et al. 2013) and present day (Mitchell and Jones, 2005) climate data seemed to capture relatively well the main spatial patterns of both glacial and interglacial climate in Europe.

2. In **Paper IV**, the modelled human population size in Europe between 30 and 13 kyr ago decreased with cooling climate and increased with warming climate. During the decrease in solar insolation between 30 and 23 kyr ago, the European mean of the mean temperature of the coldest month decreased about 8 degrees in 7,000 years. In this time period the modelled European human population size decreased from 330,000 to 130,000. Further, due to increased solar insolation, during the very end of the last glacial, between 19 and 13 kyr ago, the simulated European mean of the mean temperature of the coldest month rose about 12 degrees, and the modelled European human population size increased from 150,000 to 400,000.

3. In **Paper II**, the trend of the climate based forest fire risk, the FWI (Van Wagner, 1987), over Europe during 1960-2012, was investigated. The FWI was found to have risen in Eastern and Southern Europe during 1980-2012, whereas in Western and Northern Europe no statistically significant trends were found.

4. In **Paper I**, the expected changes in the mean and extreme winds in Northern Europe under global warming were investigated by 9 GCMs under SRES A2, B1 and A1B scenarios. The results showed that changes in the mean and extreme winds are going to be small during the ongoing century: in parts of northwestern Russia and southern Baltic Sea the winds might increase by 2-4% and over the Norwegian Sea the winds might decrease by 2-8%.

5. In **Paper V**, the statistical relationship of the QBO and the surface AO index was used in post-processing 3-6 weeks surface temperature forecasts in Northern Europe and the skill scores of these forecasts were thereby slightly improved. As the S2S database ERF models have shown better prediction skills for the stratosphere than for the troposphere (Domeisen 2019a, 2019b), and as the stratosphere has important precursors for the tropospheric weather (Kidson et al. 2015), and as in the future there might be changes in the mean stratospheric flow: I conclude that improving the modelling of the stratosphere in both climate and weather simulations seems utmost useful.

REFERENCES

Alexander, L. V., Uotila, P., and N. Nicholls, N.: Influence of sea surface temperature variability on global temperature and precipitation extremes, J. Geophys. Res.,114, 2009.

Ammann, C. M, Joos, F., Schimel, D. S., Otto-Bliesner, B. L., Tomas, R. A.: Solar influence on climate during the past millennium: Results from transient simulations with the NCAR Climate System Model, Proceedings of the National Academy of Sciences, 104, 3713-3718, 2007.

Anchukaitis, K.J., Buckley, B. M., Cook, E. R., Cook, B. I., D'Arrigo, R. D., and Ammann, C. M.: The influence of volcanic eruptions on the climate of the Asian monsoon region. Geophys. Res. Lett., 37, 2010.

Antonsson, K., Brooks, S. J., Seppä, H., Telford, R. J., and Birks, H. J. B.: Quantitative palaeotemperature records inferred from fossil pollen and chironomid assemblages from Lake Gilltjärnen, northern central Sweden, J. Quat. Sci., 21, 831–841, 2006.

Ayarzagüena, B. et al.: No robust evidence of future changes in major stratospheric sudden warmings: a multi-model assessment from CCMI, Atmos. Chem. Phys., 18, 11277–11287, 2018.

Baldwin, M. P., and Dunkerton, T.J.: Propagation of the Arctic Oscillation from the stratosphere to the troposphere, J. Geophys. Res, 104, D24, 30937-30946, 1999.

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., et al.: The quasi-biennial oscillation, Rev. Geophys., 39(2), 179–229, 2001.

Bedia, J., Herrera, S., Gutiérrez, J. M., Zavala, G., Urbieta, I. R., and Moreno, J. M.: Sensitivity of fire weather index to different reanalysis products in the Iberian Peninsula, Nat. Hazards Earth Syst. Sci., 12, 699–708, 2012.

Berger, A. & Loutre, M.F.: Insolation values for the climate of the last 10 million years. Quaternary Science Reviews, 10, 297–317, 1991.

Binford, L. R.: Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Ethnographic and Environmental Data Sets (Univ of California Press, Berkeley), 2001.

Bowman, D., Balch, K., Artaxo, P., Bond, W., Carlson, J., Cochrane, M., D'Antonio, C., DeFries, R., Doyle, J., Harrison, S., Johnston F., Keeley, J., Krawchuk, M., Kull, C., Marston, J., Moritz, M., Prentice, C., Roos, C., Scott, A., Swetnam, T., Der Werf, G., and Pyne, S.: Fire in the Earth System. Science, 324, 481–484, 2009.

Brady, E., Otto-Bliesner, B., Kay, J., and Rosenbloom, N.: Sensitivity to Glacial Forcing in the CCSM4, J. Climate, 36, 1901–1925, 2013.

Brandefelt, J. and Otto-Bliesner, B.: Equilibration and variability in a Last Glacial Maximum Climate simulation with CCSM3, Geophys. Res. Lett., 36, L19712, doi:10.1029/2009GL040364, 2009.

Buizza, R. and Leutbecher, M.: The forecast skill horizon, Q. J. R. Meteorol. Soc., 141, 3366–3382, doi:10.1002/qj.2619, 2015.

Butler, A. H., Seidel, D. J., Hardiman, S. C., Butchart, N., Birner, T., and Match, A.: Defining sudden stratospheric warmings, Bull. American Meteor. Soc., 96, 1913-1928, doi: http://dx.doi.org/10.1175/BAMS-D-13-00173.1, 2015.

Calov, R., Ganopolski, A., Claussen, M., Petoukhov, V. & Greve, R.: Transient simulation of the last glacial inception. Part I: glacial inception as a bifurcation in the climate system. Climate Dynamics, 24, 545–561, 2005.

Chambers, J. M., Cleveland, W. S., Kleiner, B, and Tukey, P.A.: Graphical Methods for Data Analysis, The Wadsworth statistics/probability series. Wadsworth and Brooks/Cole, Pacific Grove, CA, 1983.

Cheng, M. and Qi, Y.: Frontal rainfall rate distribution and some conclusions on the threshold method, J. App. Meteorol., 41, 1128–1139, 2002.

Claussen, M., Mysak, L., Weaver, A., Crucifix, M., Fichefet, T., Loutre, M.F., Weber, S., Alcamo, J., Alexeev, V., Berger, A., Calov, R., Ganopolski, A., Goosse, H., Lohmann, G., Lunkeit, F., Mokhov, I., Petoukhov, V., Stone, P. & Wang, Z.: Earth system Models of Intermediate Complexity: Closing the gap in the spectrum of climate system models, Climate Dynamics, 18, 579–586, 2002.

Cohen, J., Screen, J., Furtado, J. et al. Recent Arctic amplification and extreme mid-latitude weather. Nature Geosci 7, 627–637, 2014.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag: London, 242 pp, 2001.

Crucifix, M., Loutre, M.F., and Berger, A.: The climate response to the astronomical forcing. Space reviews, 125, 213–226, 2006.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Hólm, E., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A., Monge-Sanz, B. M., Morcrette, J., Park, B., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553–597, 2011.

Domeisen, D.I.V., Butler, A. H., Charlton-Perez, A. J., Ayarzagüena, B., Baldwin, M. P., Dunn-Sigouin, E., Furtado, J. C., Garfinkel, C. I., Hitchcock, P., Karpechko, A.Yu., Kim, H., Knight, J., Lang, A. L., Lim, E.-P. Marshall, A., Roff, G., Schwartz, C., Simpson, I. R., Son, S.-W., Taguchi, M.: The role of the stratosphere in subseasonal to seasonal prediction. Part I: Predictability of the stratosphere, Journal of Geophysical Research – Atmospheres, 2019JD030920–38. http://doi.org/10.1029/2019JD030920, 2019a.

Domeisen, D.I.V., Butler, A. H., Charlton-Perez, A. J., Ayarzagüena, B., Baldwin, M. P., Dunn-Sigouin, E., Furtado, J. C., Garfinkel, C. I., Hitchcock, P., Karpechko, A.Yu., Kim, H., Knight, J., Lang, A. L., Lim, E.-P. Marshall, A., Roff, G., Schwartz, C., Simpson, I. R., Son, S.-W., Taguchi, M.: The role of the stratosphere in subseasonal to seasonal prediction. Part II: Predictability arising

from stratosphere - troposphere coupling, Journal of Geophysical Research – Atmospheres, 2019JD030923. http://doi.org/10.1029/2019JD030923, 2019b.

Ehlers, J. and Gibbard, P. L. (Eds.): Quaternary glaciations – extent and chronology. Part I: Europe. Elsevier, 475 pp, 2004.

EPICA community members 2004. Eight glacial cycles from an Antarctic ice core. Nature, 429(6992), 623–628, 2004.

Feser, F., Barcikowska, M., Krueger, O., Schenk, F., Weissea, R., and Xiae, L.: Storminess over the North Atlantic and northwestern Europe – A review, Q. J. R. Meteorol. Soc. 141: 350–382, January 2015 B DOI:10.1002/qj.2364, 2015.

Finnish Forest Research Institute: Finnish Statistical Yearbook of Forestry. Finnish Forest Research Institute, 472 pp., Helsinki, Finland, 2010 (in Finnish with summary and captions in English).

Frank, D., Esper, J., Raible, C. et al. Ensemble reconstruction constraints on the global carbon cycle sensitivity to climate. Nature 463, 527–530, 2010.

Ganopolski, A., Petoukhov, V., Rahmstorf, S., Brovkin, V., Claussen, M., Eliseev, A. & Kubatzki, C.: CLIMBER-2: a climate system model of intermediate complexity. Part II: model sensitivity. Climate Dynamics, 17, 735–751, 2001.

Ganopolski, A., Calov, R. & Claussen, M.: Simulation of the last glacial cycle with a coupled climate ice-sheet model of intermediate complexity. Climate of the Past, 6, 229–244, 2010.

Garfinkel, C. I., Schwartz, C., Domeisen, D. I. P., Son, S-W., Butler, A. H., and White, I. P.: Extratropical stratospheric predictability from the Quasi-Biennial Oscillation in Subseasonal forecast models, J. Geophys. Res: Atmospheres, doi: 10.1029/2018JD028724, 2018.

Gray, L. J., Anstey, J. A., Kawatani, Y., Lu, H., Osprey, S., and Schenzinger, V.: Surface impacts of the Quasi Biennial Oscillation, Atmos. Chem. Phys., 18, 8227–8247, https://doi.org/10.5194/acp-18-8227-2018, 2018.

Gregow H, Peltola H, Laapas M, Saku S, Venäläinen A.: Combined occurrence of wind, snow loading and soil frost with implications for risks to forestry in Finland under the current and changing climatic conditions. Silva Fennica 45(1): 35–54, 2011b.

Greve, R.: Application of a polythermal three-dimensional ice sheet model to the Greenland ice sheet: Response to steady-state and transient climate scenarios. Journal of Climate, 10, 901–918, 1997.

Haigh, J. D.: The Impact of Solar Variability on Climate, Science, 272, 981-984, 1996.

Harvey, L. D. D.: Global Warming: The hard science. Pearson Education Limited, Harlow, 336 pp. ISBN 0582-381167-3, 2000.

Hastie, T. and Tibshirani, R.: Generalized Additive Models, monographs on statistics and applied probability, vol. 43, Chapman & Hall, New York, 1990.

Heikkilä, M. and Seppä, H.: A 11,000yr palaeotemperature reconstruction from the southern boreal zone in Finland, Quat. Sci. Rev., 22, 541–554, 2003.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems. Wea. Forecasting, 15, 559–570, doi:10.1175/1520-0434(2000)015,0559: DOTCRP.2.0.CO;2, 2000.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol., 25(15):1965–1978, 2005.

Holton, J. R. and Tan, H. C.: The influence of the equatorial quasi-biennial oscillation on the global circulation at 50mb, J. Atmos. Sci., 37, 2200–2208, 1980.

Holton, J. R. and Tan, H. C.: The quasi-biennial oscillation in the Northern Hemisphere lower stratosphere, J. Meteor. Soc. Japan, 60, 140–148, 1982.

IPCC: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 881pp, 2001.

IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp, 2007.

IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, 2013.

IPCC: Summary for policymakers. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1-32, 2014.

IPCC: “Summary for Policymakers. Global Warming of 1.5°C.” In Global Warming of 1.5°C. An IPCC Special Report on the Impacts of Global Warming of 1.5°C above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, eds. V. Masson-Delmotte et al. World Meteorological Organization, Geneva, Switzerland, 32 pp, 2018.

Joughin, I., and Alley, R. Stability of the West Antarctic ice sheet in a warming world. Nature Geosci 4, 506–513, 2011.

Katz, R. W., Bruch, G.S., Parlange, M. B.: Statistics of extremes: modelling ecological disturbances. Ecology 86: 1124–1134, 2005.

Kawatani, Y., Hamilton, K., Watanabe, S.: The quasi-biennial oscillation in a double CO2 climate. J Atmos Sci 68(2):265–283, 2011.

Keeling, C.D., and T.P. Whorf: Atmospheric CO_2 records from sites in the SiO air sampling network. In: Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, 2005.

Kendall, M. G.: Rank correlation methods, Griffin, London, 1975.

Kidston, J.; Scaife, A. A.; Hardiman, S. C.; Mitchell, D. M.; Butchart, N.; Baldwin, M. P., and Gray, L. J.: Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. Nature Geoscience, 8(6), 433–440, 2015.

Kjellström, E., Brandefelt, J., Näslund, J. O., Smith, B., Strandberg, G., Voelker, A. H. L., and Wohlfarth, B.: Simulated climate conditions in Fennoscandia during a MIS 3 stadial, Boreas, 39, 436– 456, 2010.

Koutsias, N., Xanthopoulos, G., Founda, D., Xystrakis, F., Nioti, F., Pleniou, M., Mallinis, G., and Arianoutsou, A.: On the relationship between forest fires and weather conditions in Greece from long-term national observations (1894–2010), Int. J. Wildland Fire, 22, 493–507, 2013.

Lehtonen, I., Ruosteenoja, K., Venäläinen, A., and Gregow, H.: The projected 21st century forest fire risk in Finland under different greenhouse gas scenarios. Boreal Environment Research, 19: 127–139, 2014.

Lehtonen, I., Venäläinen, A., Kämäräinen, M., Peltola, H., and Gregow, H.: Risk of large-scale fires in boreal forests of Finlandunder changing climate, Nat. Hazards Earth Syst. Sci., 16, 239–253, 2016.

Lenssen, N., Schmidt, G., Hansen, J., Menne, M., Persin, A., Ruedy, R., and Zyss, D.: Improvements in the GISTEMP uncertainty model. J. Geophys. Res. Atmos., 124, no. 12, 6307-6326, 2019.

Limpasuvan, V., Hartmann, D. L., Thompson, D. W. J., Jeev, K., and Yung, Y. L.: Stratosphere-troposphere evolution during polar vortex intensification. J. Geophys. Res., 110, D24101, doi: 10.1029/2005JD006302, 2005.

MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., van Ommen, T., Smith, A., and Elkins, J.: The Law Dome CO_2, CH_4 and N_2O ice core records extended to 2000 years BP. Geophys. Res. Lett., 33, L14810, doi:10.1029/2006GL026152, 2006.

Mann, H. B.: Nonparametric tests against trend, Econometrica, Econometrica, 13, 245–259, 1945.

Manzini, E., et al.: Northern winter climate change: Assessment of uncertainty in CMIP5 projections related to stratosphere-troposphere coupling, J. Geophys. Res. Atmos., 119, 7979–7998, 2014.

Masson-Delmotte, V., M. Schulz, A. Abe-Ouchi, J. Beer, A. Ganopolski, J.F. González Rouco, E. Jansen, K. Lambeck, J. Luterbacher, T. Naish, T. Osborn, B. Otto-Bliesner, T. Quinn, R. Ramesh, M. Rojas, X. Shao and A. Timmermann: Information from Paleoclimate Archives. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, US, 2013.

Meehl, G.A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J, Taylor, K. E.: 2007. The WCRP CMIP3 multimodel dataset: a new era in climate change research. Bulletin of American Society 88: 1383–1394, 2007.

Mikkonen, S., Laine, M., Mäkelä, H. M., Gregow, H., Tuomenvirta, H., Lahtinen, M. & Laaksonen, A. Trends in the average temperature in Finland, 1847–2013. Stochastic Environmental Research and Risk Assessment, 2015.

Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, 2005.

Monnin, E., E.J. Steig, U. Siegenthaler, K. Kawamura, J. Schwander, B. Stauffer, T.F. Stocker, D.L. Morse, J.-M. Barnola, B. Bellier, D. Raynaud, and H. Fischer: Evidence for substantial accumulation rate variability in Antarctica during the Holocene, through synchronization of CO_2 in the Taylor Dome, Dome C and DML ice cores. Earth and Planetary Science Letters, 224, 45-54, 2004.

Mäkelä, H. M., Venäläinen, A., Jylhä, K., Lehtonen, I., and Gregow, H.: Probablistic projections of climatological forest fire danger in Finland. Climate Research, 60, 73-85, 2014.

Nakicénovic N, Alcamo J, Davis G, de Vries B, Fenhann J, Gaffin S, Gregory K, Gruebler A, Jung TY, Kram T, La Rovere EL, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Riahi K, Roehrl A, Rogner H-H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z.: IPCC Special Report on Emissions Scenarios. Cambridge University Press: Cambridge and New York, 599 pp, 2000.

Naujokat, B.: An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. J. Atmos. Sci., 43, 1873–1877. 1986.

Otto-Bliesner, B.L., Brady, E.C., Clauzet, G., Tomas, R., Levis, S., and Kothavala, Z.: Last glacial maximum and Holocene climate in CCSM3, J. Climate, 19, 2526–2544, 2006.

Pearson R. G., and Dawson T. P.: Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? Glob Ecol Biogeogr 12(5):361–371, 2003.

Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J. M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, V. M., Legrand, M., Lipenkov, V.Y., Lorius, C., Pepin, L., Ritz, C., Saltzman, E. & Stievenard, M.: Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature, 399, 429–436, 1999.

Petoukhov, V., Ganopolski A., Brovkin V., Claussen M., Eliseev A., Kubatzki & Rahmstorf S.: CLIMBER-2: A climate system model of intermediate complexity. Part I: Model description and performance for present climate. Climate Dynamics, 16, 1–17, 2000.

Petoukhov, V., Claussen, M., Berger, A., Crucifix, M., Eby, M., Eliseev, A. V., Fichefet, T., Ganopolski, A., Goosse, H., Kamonkovich, I., Mokhov, I. I., Montoya, M., Mysak, L. A., Sokolov, A., Stone, P., Wang, Z., and Weaver, A. J.: EMIC Intercomparison Project (EMIP-CO2): comparative analysis of EMIC simulations of climate, and of equilibrium and transient responses to atmospheric CO2 doubling, Clim. Dynam., 25, 363– 385, 2005.

Rienecker et al. MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, Journal of Climate, 24, 3624-3648, 2011.

Ruosteenoja, K., Vihma, T., and Venäläinen, A.: Projected Changes in European and North Atlantic Seasonal Wind Climate Derived from CMIP5 Simulations, Journal of Climate, 19, 6467-6490, 2019.

Scaife, A. A., et al.: Predictability of the quasi-biennial oscillation and its northern winter teleconnection on seasonal to decadal timescales, Geophys. Res. Lett., 41, 1752–1758, doi:10.1002/2013GL059160, 2014.

Schirber, S., Manzini, E., Krismer, T. & Giorgetta, M.: The quasi-biennial oscillation in a warmer climate: Sensitivity to different gravity wave parameterizations, Clim. Dyn. 45, 825-836, 2015.

Schoeberl, M. R.: Stratospheric warmings: Observations and theory. Rev. Geophys., 16, 521–538, 1978.

Shaffer, G. & Lambert, F.: Glacial extremes from dust−climate feedbacks, Proceedings of the National Academy of Sciences, 115 (9) 2026-2031, 2018.

Smith, R. S. and Gregory, J.: The last glacial cycle: transient simulations with an AOGCM. Clim. Dynam., 38, 1545–1559, 2012.

Tallavaara, M., Eronen, J. T., Miska Luoto, M.: Drivers of the hunter-gatherer population density, Proceedings of the National Academy of Sciences, 115 (6) 1232-1237, 2018.

Thompson, D. W. J., Baldwin, M. P. and Wallace J. M.: Stratospheric connection to Northern Hemisphere wintertime weather: implications for prediction. J. Clim. 15, 1421–1428, 2002.

Tomassini, L., Gerber, E. P., Baldwin, M. P., Bunzel, F. and Giorgetta, M.: The role of stratosphere troposphere coupling in the occurrence of extreme winter cold spells over northern Europe. J. Adv. Model. Earth Syst., 4, M00A03, 2012.

Uppala, S., Kallberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R., Andersson, E., Arpe, K., Balmaseda, M., Beljaars, A., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B., Isaksen, L., Janssen, P., Jenne, R., McNally, A., Mahfouf, J., Morcrette, J., Rayner, N., Saunders, R., Simon, P., Sterl, A., Trenberth, K., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Q. J. R. Meteorol. Soc., 131, 2961–3012, 2005.

van Vuuren, D. P., and Coauthors: The representative concentration pathways: an overview. Climatic Change, 109, 5–31, 2011.

Van Wagner, C.: Development and structure of the Canadian Forest Fire Weather Index System. Canadian Forestry Service, Ottawa, Ontario, Forestry Technical Report, 35, 37 pp., 1987.

Vermeersch P. M.: European population changes during Marine Isotope Stages 2 and 3. Quat Int 137(1):77–85, 2005.

Vitart F.: Evolution of ECMWF sub-seasonal forecast skill scores. Q. J. R. Meteorol. Soc., 140, 1889–1899, doi: 10.1002/qj.2256, 2014.

Vrac, M., Marbaix, P., Paillard, D. & Naveau, P.: Non-linear statistical downscaling of present and LGM precipitation and temperature over Europe. Climate of the Past, 3, 669-682, 2007.

Watson, P. A., and L. J. Gray: How Does the Quasi-Biennial Oscillation Affect the Stratospheric Polar Vortex? J. Atmos. Sci.,71, 391–409, doi: 10.1175/JAS-D-13-096.1, 2014.

Wessel, P. & Müller, R. D.: Plate Tectonics, Treatise on Geophysics, 6, Elsevier, pp. 45-93, 2015.

Wood, S.N.: Generalized additive Models: An Introduction with R. CRC/Chapman & Hall, 2006.

Paper I

© 2011 Royal Meteorological Society

Reprinted, with permission from

International Journal of Climatology, 32, 1834–1846,

doi: 10.1002/joc.2398

INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **32**: 1834–1846 (2011)
Published online 2 August 2011 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.2398

RMetS
Royal Meteorological Society

Changes in the mean and extreme geostrophic wind speeds in Northern Europe until 2100 based on nine global climate models

Hilppa Gregow,* Kimmo Ruosteenoja, Natalia Pimenoff and Kirsti Jylhä
Finnish Meteorological Institute, Helsinki, Finland

ABSTRACT: This study aims at analyzing the mean and extreme geostrophic wind speeds in Northern Europe. The analyses are based on nine global climate models and the Special Report on Emission Scenarios (SRES) A1B, A2 and B1 scenarios. The time frames studied consist of the baseline 1971–2000 and the future periods 2046–2065 and 2081–2100. The SRES scenarios are considered both separately and combined. The extremes are calculated for the September–April period for various return periods. The analysis is done by applying the program R and the Generalized Extreme Value -methodology.

All projections indicate that both the mean and extreme geostrophic wind speeds will increase in the southern and eastern parts of Northern Europe and decrease over the Norwegian Sea in September–April. The change over the ocean is pronounced already in 2046–2065, over the continents in 2081–2100. For the model mean, the smallest change (2–6%) was projected under the B1 and the largest (4–10%) under the A1B and A2 scenarios. However, spread among the individual global circulation models (GCMs) was fairly large.

The ratios between the return level estimates for various return periods and the annual maximum wind speeds were found nearly homogeneously independent of the time frame studied. For the baseline and future periods, the extreme winds occurring once in 10 or 50 years were $13\% \pm 2\%$ and $22\% \pm 5\%$ stronger than the mean annual maxima, respectively. The present findings serve as support for risk assessment such as required when planning the forest management practices. Copyright © 2011 Royal Meteorological Society

KEY WORDS geostrophic wind; global climate models; emission scenarios; wind extremes; uncertainty; sensitivity

Received 11 November 2010; Revised 23 May 2011; Accepted 21 June 2011

1. Introduction

The socio-economic impacts of the violent extreme winds caused by extratropical cyclones have been anomalously large in Europe during the last two decades. For example, in the Decembers of 1990 and 1999, respectively, a total of 100 and 175 Mm^3 of timber blew down in winter storms throughout Europe (Ulbrich *et al*., 2001; Dobbertin, 2002; Schönenberger, 2002; Brüdl and Rickli, 2002). In January 2005, about 70 Mm^3 of timber was damaged in Sweden (Alexandersson, 2005; Bengtsson and Nilsson, 2007) and in January 2007, about 45 Mm^3 in Central Europe (Fink *et al*., 2009). In Finland, the most destructive recent wind damages occurred in November 2001 and in July–August 2010. These storms caused a loss of 7.3 and 8 Mm^3 of timber, respectively. The economic influence of the damaging winds as well as the secondary damages caused, for instance, by the insect attacks on the remaining trees (Nykänen *et al*., 1997; Valinger and Fridman, 1997; Schönenberger, 2002; Jönsson *et al*., 2009; Jönsson and Bärring, 2010) are particularly harmful in the managed forests. This is because the value of harvested timber reduces due to the sudden excess of supply, at the same time the unscheduled and less optimal harvesting procedures increase the costs (Peltola *et al*., 2010).

Because of the recurrent damage to infrastructure and forests by windstorms (Ulbrich *et al*., 2001; Dobbertin, 2002; Wernli *et al*., 2002; Pellikka and Järvenpää, 2003; Alexandersson, 2005; Pinto *et al*., 2007), the cyclone frequency and intensity in Europe have gained a lot of emphasis in research. For example, Bärring and Fortuniak (2009) found pronounced interdecadal variability in cyclone activity but no significant long-term trend in southern Scandinavian storminess between 1780 and 2005. Correspondingly, Hanna *et al*. (2008) found no sign of enhancement in storminess associated with climatic change when investigating the Northern European and North Atlantic surface pressure variability measurements since the 1830s. Additionally, according to Bärring and Fortuniak (2009) and Matulla *et al*. (2007), in northern and central parts of Europe the links of storminess to the North Atlantic Oscillation (NAO) vary depending on the region and the time periods.

* Correspondence to: Hilppa Gregow, Finnish Meteorological Institute, P.O. Box 503, FI-00101 Helsinki, Finland.
E-mail: hilppa.gregow@fmi.fi

Copyright © 2011 Royal Meteorological Society

In Northern Europe, strong winds are most frequent in the cold season (September–April) because then the spatial variations of temperature are highest (Lau, 1988, Chang *et al*., 2002) and the conversion of available potential energy to kinetic energy largest. According to Donohoe and Battisti (2009), also of great importance is the global stationary wave pattern that allows the Pacific and Atlantic storm tracks to be connected. If the Pacific and Atlantic cyclone tracks are disconnected, the storms can develop rapidly but the intensities are weaker than if the cyclone tracks are connected. This is independent of the abundance of available potential energy. In line with Donohoe and Battisti (2009), in the investigations of Ulbrich *et al*. (2008), an increase in the number of the deepest cyclones takes place in the northern North Pacific in tandem with the eastern Northern Atlantic as the global warming proceeds. Also Bengtsson *et al*. (2009), who investigated winds at the 925 hPa level with the ECHAM5 T213 model (63 km resolution), noticed that the highest wind speeds in winter (December–February) occur predominantly over the mid-latitude areas of the Atlantic and Pacific Oceans.

To provide valid information for decision support, wind speed projections based on various models and greenhouse gas (GHG) scenarios should be analyzed in great detail, considering also the associated uncertainties. In this work, we study the modelled responses in the mean and extreme geostrophic wind speeds (hereafter mean and extreme V_g) to GHG forcing. The V_g speeds are considered rather than the true surface wind speeds because the former are less affected by model parameterization than the latter (Rõõm 1998, Zilitinkevich *et al*. 2002, Schulz, 2008). The analysis concentrates on Northern Europe, especially Finland and the Baltic Sea. The calculations are made using data from simulations performed with nine global circulation models (GCMs) employing the Special Report on Emission Scenarios (SRES, Nakićenović *et al*., 2000) A1B, A2 and B1 scenarios. The periods to be compared represent the baseline climate 1971–2000 and projections for 2046–2065 and 2081–2100; for the future climate, model data at daily level is only available for these discrete 20-year periods. The surface geostrophic wind speeds are derived from the model-simulated daily mean sea level pressure and temperature fields. The analyses focus on the high-wind season from September to April (SepApr).

The main goal of this paper is to explore the projected changes in the mean and extreme geostrophic wind speeds. The simulated baseline period temporal mean geostrophic wind speeds of nine GCMs are first compared to the observational estimates derived from the ERA-40 dataset (Uppala *et al*., 2005). Then the methods used for the calculation of the projected changes in the mean, maximum and extreme wind speed are described. A comprehensive analysis of extreme wind speeds for the various return periods is performed for all grid points using the Generalized Extreme Value (GEV) theory and the block maxima approach. For comparison, a simplified method that utilizes the close relationship between the n-year return level estimates and the average annual maximum is developed. This simple method is applicable to daily practices, for instance, in climate centres, insurance companies or other service providers, where service based on climatic statistics is given based on the continuously updated datasets. The uncertainty in the projected changes is explored by finding out the difference among the various GHG scenarios, individual model results as well as the two methods applied in the extreme analyses.

2. Material and methods

2.1. Climate model data

In this work, altogether nine GCMs (Table I) were used to study the changes in the surface geostrophic wind speeds in Northern Europe (50°N, 10 °W → 80 °N, 40 °E). Three periods, i.e., years 1971–2000, 2046–2065 and 2081–2100 were examined. The nine models utilized in this work are a subset of the 23 models used in the Intergovernmental Panel on Climate Change 4th Assessment Report in 2007. These GCMs represent separate institutes, and only models with spatial resolution of about 300 km (T42) or denser were included in the analysis. The model-simulated sea level pressure and the temperature data were downloaded from the Coupled Model Intergovernmental Project 3 (CMIP3)

Table I. The outlines for the climate models employed in this work (see details for models from IPCC, 2007, Table VIII.1).

Model ID	Institute	Resolution	Scenarios
BCCR-BCM2.0[a]	Bjerknes Centre for Climate Research	T63 (1.9° × 1.9°)	A1B,A2
CGCM3.1(T63)	Canadian Centre for Climate Modelling and Analysis	T63 (1.9° × 1.9°)	A1B,B1
CNRM-CM3	Météo-France	T63 (1.9° × 1.9°)	A1B,A2,B1
ECHAM5/MPI-OM	Max Planck Institute for Meteorology	T63 (1.9° × 1.9°)	A1B,A2,B1
GFDL-CM2.1	National Oceanic and Atmospheric Administration (NOAA)	2.0° × 2.5°	A1B,A2,B1
IPSL-CM4	Pierre Laplace Institute	2.5° × 3.75°	A1B,A2,B1
MIROC3.2(hires)	Japan Center for Climate System Research	T106 (1.1° × 1.1°)	A1B,B1
MRI-CGCM2.3.2	Japan Meteorological Research Institute	T42 (2.8° × 2.8°)	A1B,A2,B1
NCAR-CCSM3[b]	National Center for Atmospheric Research	T85 (1.4° × 1.4°)	A1B,A2,B1

[a] The data of BCCR-BCM2.0 available for years 1971–1998, 2046–2065 and 2081–2098.
[b] The data of NCAR-CCSM3 analyzed for years 1970–1999, 2045–2064 and 2080–2099.

Copyright © 2011 Royal Meteorological Society

archive (Meehl *et al.*, 2007). For the baseline period 1971–2000, one simulation for each model for the 30-year period was available. For the future, however, only 20-year periods 2046–2065 and 2081–2100 with two to three GHG simulations per model (Table I) were available.

2.2. Calculation of the geostrophic wind speed

Because the CMIP3 data archive does not provide any projections for the scalar wind velocity, we calculated the daily mean geostrophic wind vector according to Equation (1):

$$\mathbf{V}^{\mathrm{GCM}}_{\mathrm{gdaily}} = (u^{\mathrm{GCM}}_{\mathrm{gdaily}}, v^{\mathrm{GCM}}_{\mathrm{gdaily}}) = \frac{\mathrm{RT}_{\mathrm{daily}}}{f p_{\mathrm{daily}}} \hat{k} \times \nabla p_{\mathrm{daily}} \qquad (1)$$

where u_{gdaily} and v_{gdaily} denote the 24-h means for the zonal and meridional component of the wind, R the gas constant of air, T_{daily} the 24-h mean temperature, f the Coriolis parameter and p_{daily} the 24-h mean surface air pressure (SLP).

The components $u^{\mathrm{GCM}}_{\mathrm{gdaily}}$ and $v^{\mathrm{GCM}}_{\mathrm{gdaily}}$ were calculated employing the intermediate points of the native grids of the separate GCMs. This allowed us to use the smallest possible grid size in the calculation. Interpolating the component data linearly onto a common $2.5^\circ \times 2.5^\circ$ grid enabled us to build the multimodel statistics. The observational data applied for evaluation of the model data was represented on the same grid. Time series for the 24-h mean geostrophic wind vector were constructed for all the nine GCMs for the three time spans and the three SRES scenarios. The magnitude of the daily mean geostrophic wind vector can be used as a qualitative measure for the geostrophic wind velocity. This quantity systematically underestimates the daily mean scalar geostrophic wind speed, but the bias occurs both in the baseline period and future GCM-simulated climate. Therefore, by expressing the projected changes in relative terms, despite this bias we can have an idea of how the long-term mean wind conditions will be altered in the future; for the utility of this so-called delta-change approach, see Räisänen (2007) and references therein. For brevity, in this paper the magnitude of the daily mean geostrophic wind vector is simply termed the 'geostrophic wind speed'.

Using the daily geostrophic wind speed V_{gdaily} on the $2.5^\circ \times 2.5^\circ$ grid, we calculated the monthly mean V_{g}(Month). The high-wind season statistics, V_{g}(SepApr) corresponding to the mean and V_{gx}(SepApr) to the maximum wind speed of each year, were calculated from 1 September to 30 April and averaged over the number of years available in the model simulations. The calculations were done for each model for the 30-year baseline climate 1971–2000 and for each model and scenario for the two 20-year periods 2046–2065 and 2081–2100. The annual maximum wind speeds of each scenario were combined to form samples. This meant that for the baseline, we had only 30 maxima per GCM. Concerning the future periods, six of the GCMs had all three scenarios (A1B : A2 : B1) and therefore 60 maxima per sample. The other GCMs had only two scenarios either A1B : A2 or A1B : B1 meaning 40 maxima to be used in the extreme analyses. The results based on the model-wise extreme analyses were combined to form the nine-GCM A1B : A2 : B1 ensemble. The significance of the projected changes was tested by using the two tailed Student's t-test.

Table II. Grid points used in the monthly mean and extreme wind speed investigations.

Grid point name	Latitude °N	Longitude °E
Baltic Sea/SEA	55	15
Helsinki	60	25
Joensuu/FOREST	62.5	30
Jyväskylä	62.5	25
Kajaani	65	27.5
Kauhava	62.5	22.5
Rovaniemi	65	25
Sodankylä	67.5	25

The names of the points refer to synoptic weather stations close to the points. The extreme value analysis was done for the two grid points that have been provided with a specific name (SEA, FOREST).

The grid points used in the detailed investigations included one grid point in the southern Baltic Sea and seven land grid points that are located close to the wind measurement stations in Finland (Table II and the Appendix of this paper). For the simplified extreme value analysis, extreme winds on several return periods were calculated for the sea grid point and one of the land grid points. These grid points have been provided with specific names SEA and FOREST.

2.3. Assessing the quality of the GCM data

The quality of the model-inferred geostrophic winds was assessed by comparing them with the corresponding observation-based approximation. We employed the ERA-40 dataset (Uppala *et al.*, 2005) in which the pressure analyses are given at 6-h intervals. We first averaged the four consecutive SLP analyses to obtain the daily mean pressure and then calculated the vector ($u^{\mathrm{ERA}}_{\mathrm{gdaily}}$, $v^{\mathrm{ERA}}_{\mathrm{gdaily}}$) from the gradient of this pressure field (Equation (1)).

The systematic error of the modelled geostrophic wind speed for the high-wind season is depicted in the upper panel of Figure 1. The bias is generally fairly small, of the order of magnitude of 0.5 m s^{-1}, the 9-GCM mean tending to slightly underestimate the observed V_{g}. Larger errors are seen in the vicinity of mountainous areas in Central Europe, and in particular, close to Greenland. In those areas, it is evident that the distribution of V_{g} is seriously affected by reduction of the surface pressure to standard sea level.

The seasonal cycle of the modelled and observation-based geostrophic wind speeds at four grid points are compared in the lower panels of Figure 1. The observation-derived time mean geostrophic wind is weakest in summer, about 7 m s^{-1}, and strongest,

Copyright © 2011 Royal Meteorological Society

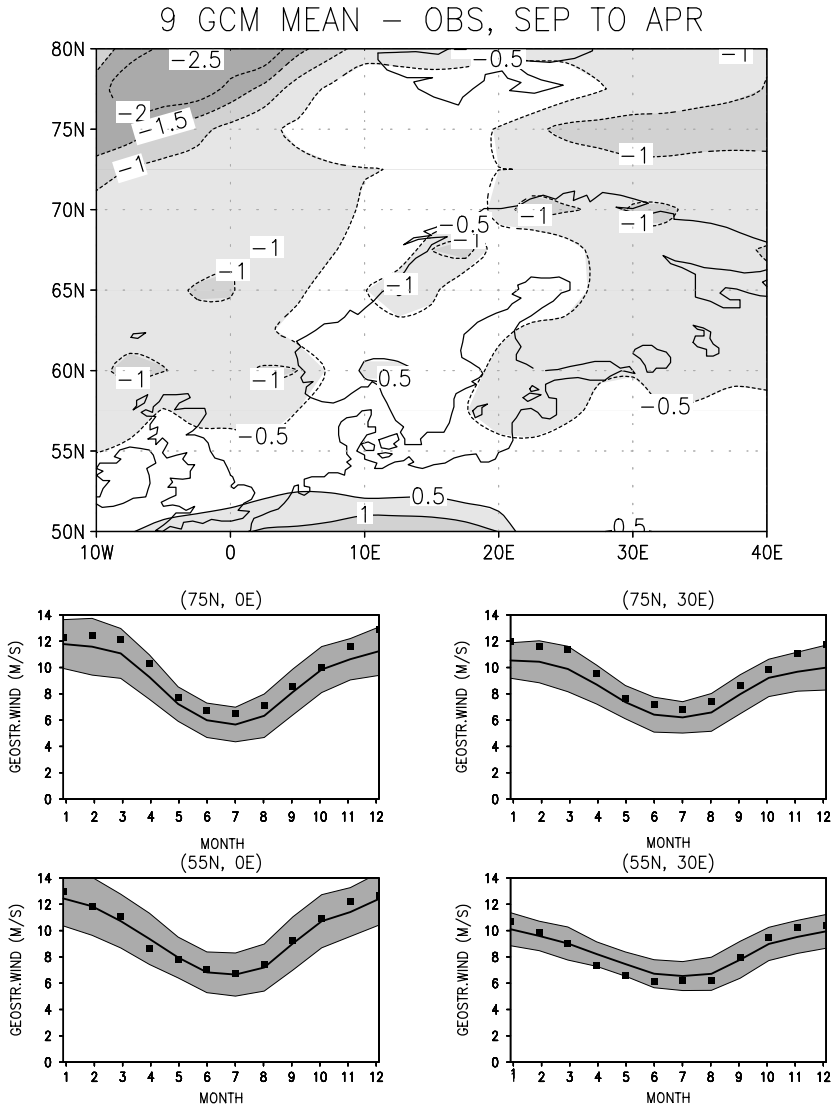

Figure 1. Map: The difference of the 9-GCM mean geostrophic wind speed for the September to April season from its observation-based counterpart derived from the ERA-40 dataset; an average for the baseline period 1971–2000, with a contour interval of 0.5 m s^{-1}. Small diagrams: The seasonal cycle of V_g at four locations. The thick line shows the multimodel mean and the shading the ± one standard deviation interval, derived from the simulations performed with the 9-GCMs. The corresponding geostrophic wind speeds calculated from the ERA-40 dataset are denoted by dots.

11–13 m s^{-1}, in winter or late autumn. The seasonal cycle is reproduced by the models reasonably well, even though the amplitude tends to be somewhat underestimated. Almost invariably, however, the observed values differ from the multimodel mean by less than one standard deviation of the modelled values.

2.4. Calculation of the extreme geostrophic wind speeds by two alternatives

2.4.1. Primary method: block maxima with the gridded dataset

The datasets of the modelled V_{gx}(SepApr) were analyzed using the GEV theory (Coles, 2001; Castillo *et al*., 2004; Wehner *et al*., 2010). We used the block maxima method by employing the program R, a free software developed by the National Center for Atmospheric Research (NCAR) (Katz *et al*., 2005) and downloadable from the internet to be used for instance for calculating extreme values of climate parameters.

The estimated probabilities of the extreme geostrophic wind speeds are expressed in terms of the return periods and the corresponding return level estimates. The return level estimates are defined as the threshold that is exceeded at any given period with the probability of $p = 1/T$, T being referred to as the return period (RP) (or waiting time) in years.

First, the GEV distributions were fitted to the sampled annual maxima using the cumulative distribution function (Equation (2)):

$$F(x;\mu,\sigma,\xi) = \begin{cases} \exp\left[-\exp\left(-\dfrac{x-\mu}{\sigma}\right)\right], \xi = 0 \\ \exp\left[-\left(1+\xi\dfrac{x-\mu}{\sigma}\right)^{-1/\xi}\right], \\ \quad \xi \neq 0,\ 1+\xi\dfrac{x-\mu}{\sigma} > 0, \end{cases} \tag{2}$$

Copyright © 2011 Royal Meteorological Society

where μ, σ and ξ represent the location, scale and shape parameters of the distribution function, respectively. The sign of the shape parameter ξ defines the type of the GEV distribution. When $\xi = 0$, the distribution is of Gumbel type (light-tailed), when $\xi < 0$, the distribution is of the bounded Weibull type and in the case of $\xi > 0$, the distribution represents the Frechet distribution with a heavy tail on the right.

After the parameters μ, σ and ξ were estimated from the sample of annual maxima, the return level estimates $V_{gx}(\mathrm{RP})$ could be determined for various return periods RP (e.g., 10 and 50 years) using Equation (3):

$$V_{gx}(\mathrm{RP}) = \begin{cases} \mu - \sigma \ln\left[-\ln\left(1-\frac{1}{\mathrm{RP}}\right)\right], \xi = 0 \\ \mu - \frac{\sigma}{\xi}\left[1-\left\{-\ln\left(1-\frac{1}{\mathrm{RP}}\right)\right\}^{-\xi}\right], \xi \neq 0. \end{cases} \quad (3)$$

The software toolkit 'extRemes' of the program R uses the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (Broyden, 1970, Fletcher, 1970, Goldfarb, 1970, Shanno, 1970) for optimization. With the software toolkit version that was suitable for the gridded datasets, the $V_{gx}(\mathrm{RP})$ could be determined systematically for the chosen return periods for all individual GCMs. After that, it was possible to form the 9-GCM ensemble for the combined A1B : A2 : B1 scenario and the 6-GCM ensemble utilizing the various scenarios.

2.4.2. *Simplified method: V_{gx} statistics and block maxima at two grid points*

Because the ratio of the $V_{gx}(\mathrm{RP})$ to the average annual maximum $V_{gx}(\mathrm{SepApr})$ proved to be rather homogeneous over the domain (not shown) another way of estimating the return level estimates was developed by using a new simplified approach. In this simplified approach, the detailed extreme analyses were conducted by using data at two grid points: 55°N, 15°E (SEA) and 62.5°N, 30°E (FOREST). The ratio of the extreme wind speeds V_{gx} (RP) to the average September–April maximum V_{gx} (SepApr) was then calculated for each model according to Equation (4):

$$\mathrm{CF}^{\mathrm{grid}}_{\mathrm{GCM}}(\mathrm{RP}) = V^{\mathrm{grid}}_{\mathrm{gxGCM}}(\mathrm{RP}) / V^{\mathrm{grid}}_{\mathrm{gx}}(\mathrm{SepApr})_{\mathrm{GCM}} \quad (4)$$

The grid point coefficients $\mathrm{CF}^{\mathrm{grid}}_{\mathrm{GCM}}$ obtained from Equation (4) were averaged over the nine GCMs for each return period to calculate coefficients CF_{71} for the period 1971–2000, CF_{46} for 2046–2065 and CF_{81} for 2081–2100. The values of the coefficients for the different periods proved to be nearly identical (Table III). Therefore, a time-independent coefficient, $\mathrm{CF}_{\mathrm{ALL}}$, was derived by averaging the coefficients determined for the three periods 1971–2000, 2046–2065 and 2081–2100. Using the model-wise time-dependent coefficients CF_{71}, CF_{46} and CF_{81}, we also assessed the spread of the changes (*cf* Section 3.3.2) between the future and the baseline periods.

Table III. Coefficients determining the ratio of the extreme geostrophic wind speeds $V_g(\mathrm{RP})$ to the mean annual maximum $V_{gx}(\mathrm{SepApr})$ (for definition, see text) as a function of the return level; an average over the two grid points (SEA and FOREST) and as a function of the return period and the time span.

RP (years)	5	10	50	100
$\mathrm{CF}_{71}(\mathrm{RP})$	1.08	1.13	1.22	1.26
$\mathrm{CF}_{46}(\mathrm{RP})$	1.08	1.13	1.22	1.26
$\mathrm{CF}_{81}(\mathrm{RP})$	1.08	1.14	1.23	1.27
$\mathrm{CF}_{\mathrm{ALL}}(\mathrm{RP})$	1.08	1.13	1.23	1.26

CF_{71} (RP) corresponds to the baseline period 1971–2000, CF_{46} (RP) to 2046–2065 and CF_{81} (RP) to 2081–2100. The average of the three coefficients $\mathrm{CF}_{\mathrm{ALL}}$ is shown on the bottom row.

On the basis of the average and the standard deviation of the coefficients CF for the nine GCMs and the three periods at the two grid points SEA and FOREST, we obtained upper and lower estimates for the values of $\mathrm{CF}_{\mathrm{ALL}}$ for each return period. The regression curves fitted to the minima, average and maxima of $\mathrm{CF}_{\mathrm{ALL}}$ are shown in Figure 2. These, plotted as a function of the return period, give the coefficients by which we can approximate the extreme $V_{gx}(\mathrm{RP})$ with uncertainty intervals on arbitrary return periods if $V_{gx}(\mathrm{SepApr})$ is known. According to Figure 2, the wind speeds that occur once in 5-, 10-, 50- and 100-years are approximately 8% ± 1%, 13% ± 2%, 23% ± 5% and 26% ± 6%, respectively, stronger than the average $V_{gx}(\mathrm{SepApr})$; i.e. the longer the return period, the larger the uncertainty in the coefficient.

2.5. Sensitivity analysis

As the main part of this work focused on the combined A1B : A2 : B1 scenario using an ensemble of nine GCMs, it was also of interest to find out how important the differences among the various GHG scenarios are. For that purpose we repeated all calculations for the mean

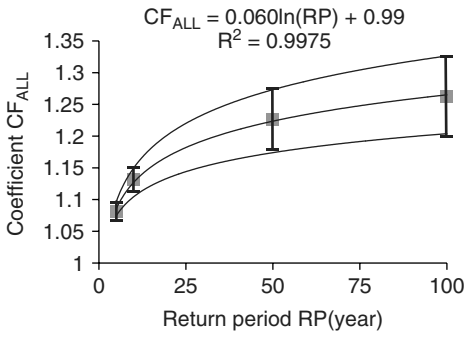

Figure 2. Coefficient CFALL (for definition see text) as a function of the return period RP. Grey boxes show the best estimate values of $\mathrm{CF}_{\mathrm{ALL}}$ for RP of 5, 10, 50 and 100 years The deviation among the model-wise estimates of the three time spans at the two grid points SEA and FOREST per return period is marked with the bars. Regression curves corresponding to the low, best and high estimate are also given.

Copyright © 2011 Royal Meteorological Society

and extreme wind speeds under the individual scenarios for a set of six GCMs that all had the three SRES scenarios available. Additionally, as the extreme analyses were performed both by using the whole gridded dataset and the simplified method, a comparison of the results of these two methods for the projected changes in the 10- and 50-year return level estimates was made.

3. Results

3.1. Mean geostrophic wind speeds

3.1.1. Projected changes

On the basis of the multimodel mean projections, the surface geostrophic wind speeds increase slightly in the southern and southeastern parts of Northern Europe and decrease over the Norwegian Sea (Figure 3). The geographical pattern of the change is qualitatively very similar for the A1B, A2 and B1 scenarios and the combined A1B : A2 : B1 ensemble scenario. This is so, although the responses to the separate scenarios are based on six rather than nine GCMs. The largest increase (>4% by 2081–2100) occurs over southern Baltic Sea and the adjacent land areas under the A1B and A2 scenarios. The smallest and statistically least significant changes are seen under the B1 forcing. The combined A1B : A2 : B1 scenario yields fairly similar changes regardless of whether a mean of six (not shown) or nine GCMs is examined. The mean wind responses for each calendar month for a number of individual locations under the A1B-scenario are reported in the Appendix. During the windiest time of the year, the monthly mean wind speeds will start to increase in the Baltic Sea already in 2046–2065 (Table AI). In Finland, increases are largest (5–7%) in November and January by 2081–2100. In November–February 2081–2100, a positive shift of 5–10% is projected to materialize in the Baltic Sea (Table AII).

3.1.2. Uncertainty

The simulated change in the monthly mean geostrophic wind speed varies from one model to another. To assess the resulting uncertainty, we fitted a normal distribution to the set of projections of changes in V_g of the individual models at the SEA and FOREST grid points, and determined the percentage points from this distribution. In winter and late autumn, the multimodel mean change in the mean V_g manifests a positive trend, up to 5–10% in a century for some months (Figure 4). In summer, the multimodel mean response is negligibly small. However, the scatter among the model simulations is quite large. At the SEA grid point, the probability for a positive change is more than 75% from October to February, whereas at the FOREST grid point this limit is only exceeded in November and January. According to the upper-estimate change (the 95th percentage point of the distribution), monthly mean wind speeds would increase by up to 15%. These increases, roughly speaking, occur in the months with the largest climatological wind speeds under the control period (Figure 1). However, as the ensemble consists of only nine GCMs, this uncertainty analysis should be considered as suggestive.

The projections also fluctuate between consecutive months (Figure 4). This holds both for the multimodel mean change and the width of the uncertainty interval. Evidently this variability is largely stochastical by its nature rather than caused by any physical grounds. For the mean of the high-wind season, the model-based probability for more intense average winds in the future is quite large, about 88% for both grid points examined in Figure 4. According to the t-test, this indicates that the deviation of the multimodel mean change from zero is confident at more than 95% level (Figure 3(b)). Conversely, the multimodel mean response is insignificantly small at both grid points (Figure 4) in the low-wind season (May–August).

3.2. Extreme geostrophic wind speeds determined from the gridded dataset with the primary method

3.2.1. Projected changes based on the combined A1B : A2 : B1 scenario

The projected changes in the extreme geostrophic wind speeds are shown in Figure 5. By the period 2046–2065, the extreme wind speeds change very little and only a small fraction of these changes are systematically significant. The pattern of change is also rather noisy. For the latter 20-year period the changes become more pronounced. The 10- and 50-year return level estimates increase by 2–4% and 2–8%, respectively, in the eastern part of the study area (20 °E–40 °E). The changes are statistically more significant for the 10-year than for the 50-year return level estimates. The decreases over the Norwegian Sea are of the order of 2–4% but only the changes in the 10-year return level estimates are statistically significant.

As far as the individual GCMs are concerned, the weakest changes by 2046–2065 were projected by IPSL and the strongest ones by CNRM (not shown). For the 20-year period 2081–2100, the weakest changes for the whole domain were produced by GFDL-CM2.1 and the largest ones by CNRM. The 10-year return level estimates produced by the individual models seemed to be rather reasonable. With the longer return periods, by contrast, anomalously high values of change appeared in the analyses around some grid points. This phenomenon occurred in the 50-year return periods in the extreme value analyses of BCCR-BCM2.0, GFDL-CM2.1 MRI-CGCM2.3.2, MIROC3.2 (hires), ECHAM5/MPI-OM and CGCM3.1 (T63). Simulated by these individual GCMs, local increases up to 20% and decreases of 16% occurred at some grid points.

3.2.2. Difference due to the GHG emission scenarios

Changes in the 10-year return level estimates of V_g as responses to the individual A1B, A2 and B1 scenarios are shown in Figure 6. As in the case of the mean V_g speed (Figure 3), the B1 forcing gives smallest and the A1B

Copyright © 2011 Royal Meteorological Society

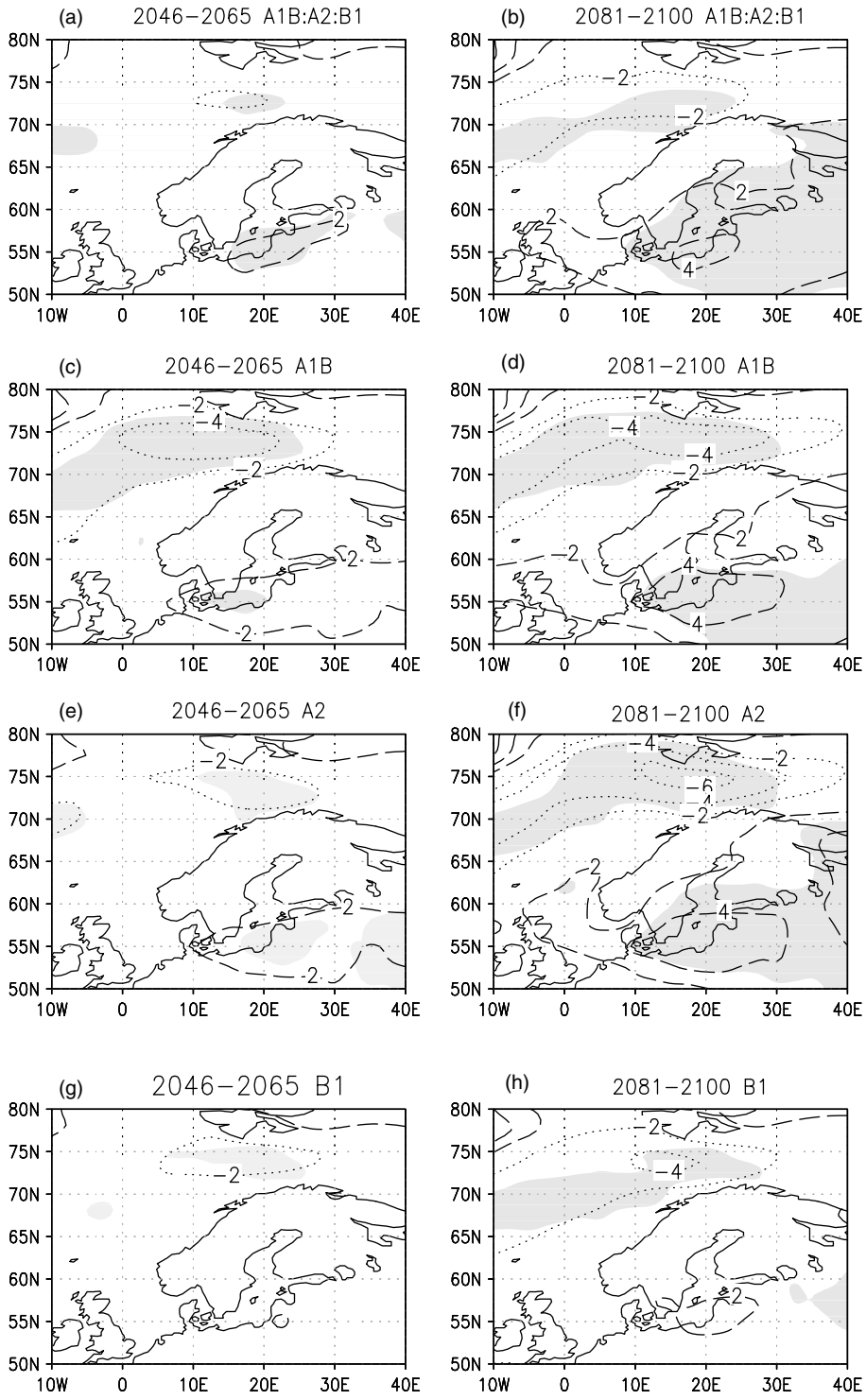

Figure 3. The multimodel mean percent changes in the September–April average geostrophic wind speeds from 1971–2000 to 2046–2065 (left) and from 1971–2000 to 2081–2100 (right). From the top (a) and (b) 9-GCM mean A1B : A2 : B1; (c) and (d) 6-GCM A1B; (e) and (f) 6-GCM A2; and (g) and (h) 6-GCM B1. Shading indicates the areas where the change is statistically significant at the 95% level. Positive changes are marked with long dashed line and negative with dots. The zero contour is omitted.

and A2 forcing largest decreases and increases. Additionally, the A1B- and A2-scenarios project fairly similar changes both in time and location (Figure 6). Positive changes of 2–6% are projected to materialize in Finland and the northwestern part of Russia. Negative changes of similar magnitude appear in the analyses over the Norwegian Sea.

3.3. Extreme wind speeds using the simplified method

3.3.1. Comparison of the two methods used for analyzing the extreme winds

The spatial distributions of the projected changes in the return level estimates of the extreme geostrophic wind speeds proved to be were qualitatively similar regardless of whether the primary (Figure 5) or the

Copyright © 2011 Royal Meteorological Society

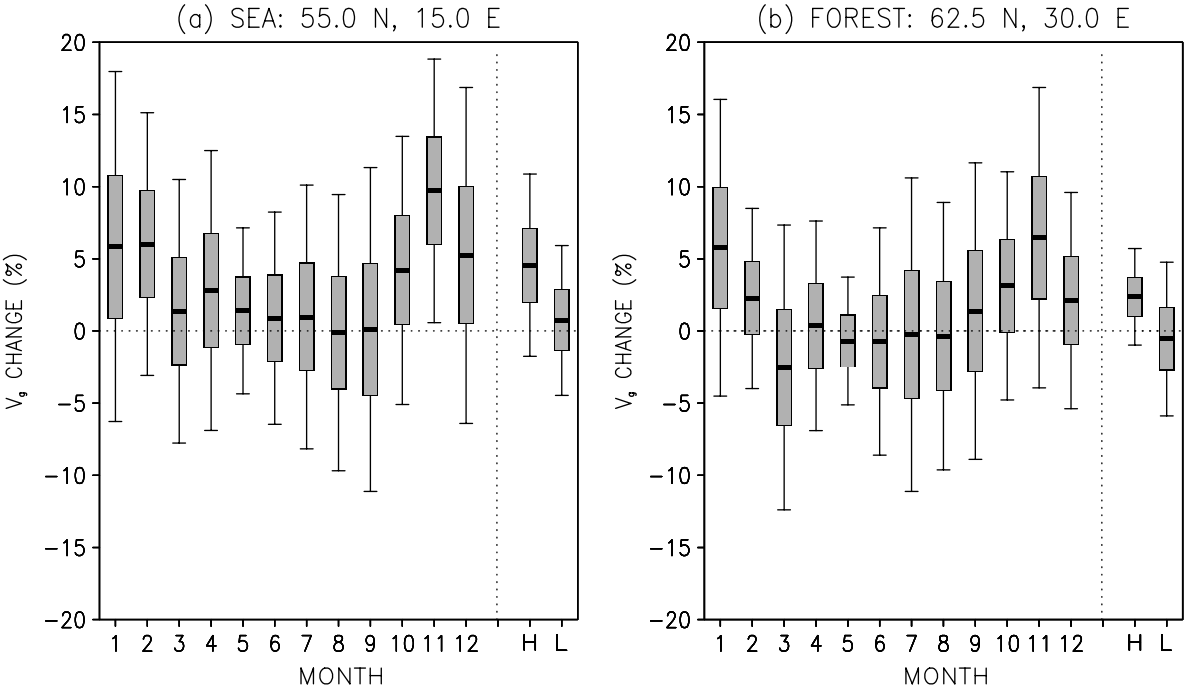

Figure 4. The seasonal cycle of the projected change in the monthly mean geostrophic wind from 1971–2000 to 2081–2100 under the A1B scenario at the two grid points SEA and FOREST: a probability distribution. The bars represent the 25–75 and whiskers the 5–95 probability interval. Numbers 1–12 refer to the calendar months. In addition to individual months, the corresponding probability intervals are depicted for the entire high- and low-wind seasons (H and L).

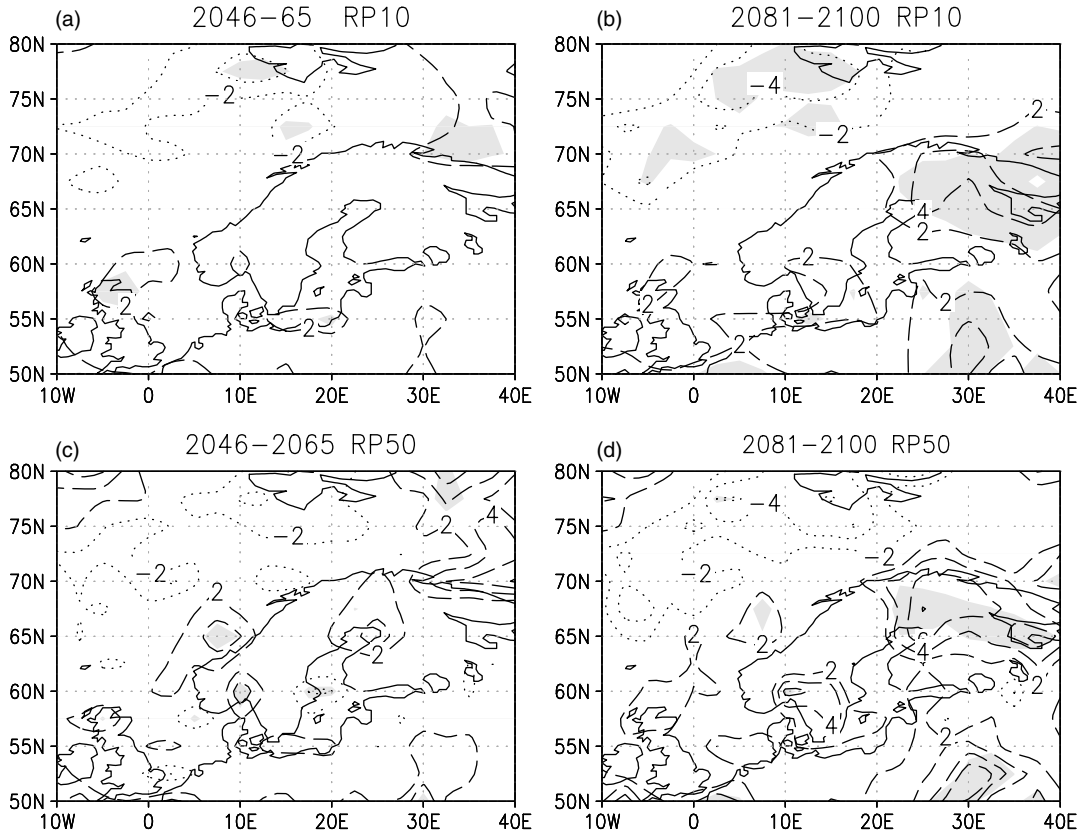

Figure 5. Changes in the return level estimates in percentages derived from the gridded dataset when comparing the periods 2046–2065 (left) and 2081–2100 (right) to 1971–2000 climate using the 9-GCM mean and the combined A1B : A2 : B1 scenario. In (a) and (b) depicted are the changes for the 10-year and in (c) and (d) for the 50-year return level estimates. Shading indicates the areas where the change is statistically significant at the 95% level. The zero contour is omitted.

Copyright © 2011 Royal Meteorological Society

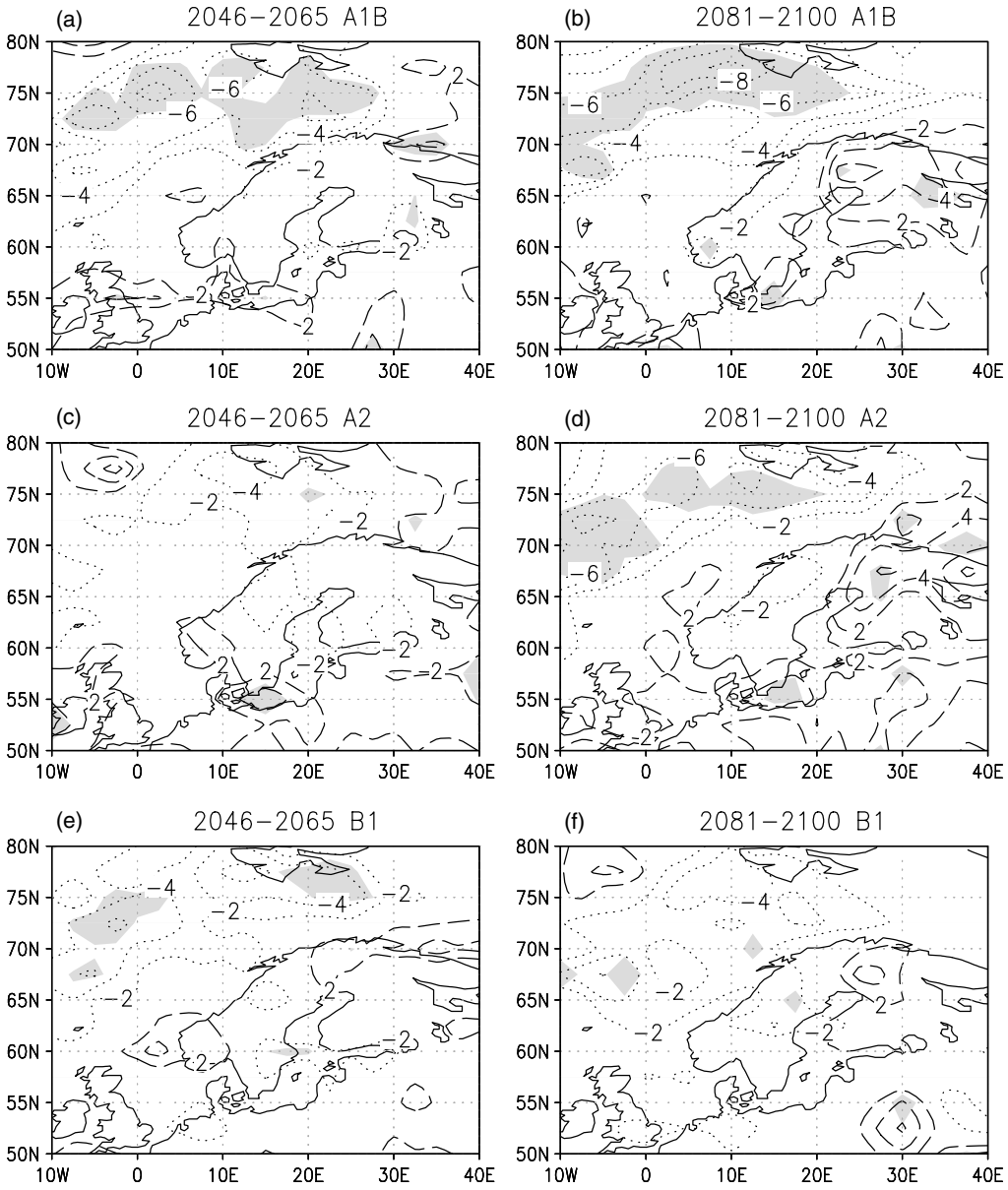

Figure 6. The changes in the 10-year return level estimates from 1971–2000 to 2046–2065 (on the left) and from 1971–2000 to 2081–2100 (on the right) for individual SRES scenarios as an average of six GCM simulations. From the top the 6-GCM mean (a) and (b) A1B; (c) and (d) A2; (e) and (f) B1. Shading indicates the areas where the change is statistically significant at the 95% level.

simplified method (Figure 7) was used. As for the gridded dataset, the simplified method gives weaker changes for 2046–2065 than for 2081–2100. Similarly, the main pattern of changes followed that described already when using complete method: an increase of 2–4% in Finland and the northwestern part of Russia and a decrease over the Norwegian Sea. One clear advantage in the simplified method is that it filters the anomalously large, evidently physically unrealistic grid point estimates that were present in six of the GCMs (see Section 3.2.1) when using the primary method.

3.3.2. Differences among the model simulations at two grid points

Using the model-wise coefficients CF_{GCM} obtained by the simplified approach and fitting a normal distribution to the set of projections based on these coefficients, we can define the percentage points for the changes of V_{gx}(RP10) and V_{gx}(RP50). These distributions are shown in Figure 8. In the extreme analyses considering the 10-year return level and the period 2046–2065, the probability of a positive change at the grid points SEA and FOREST is about 75 and 50%, respectively (Figure 8). In 2081–2100, the extreme wind speeds of the 10-year return level show, at both points, an increase at about 75% probability. For the changes of the 50-year return level, the uncertainty intervals are much broader than for the 10-year level. The interval for the period 2081–2100 ranges from about −9 to +13% at both grid points. In other words, even the sign of the change in the 50-year return level estimates of geostrophic wind speeds at these two grid points, SEA and FOREST, is highly uncertain.

Copyright © 2011 Royal Meteorological Society

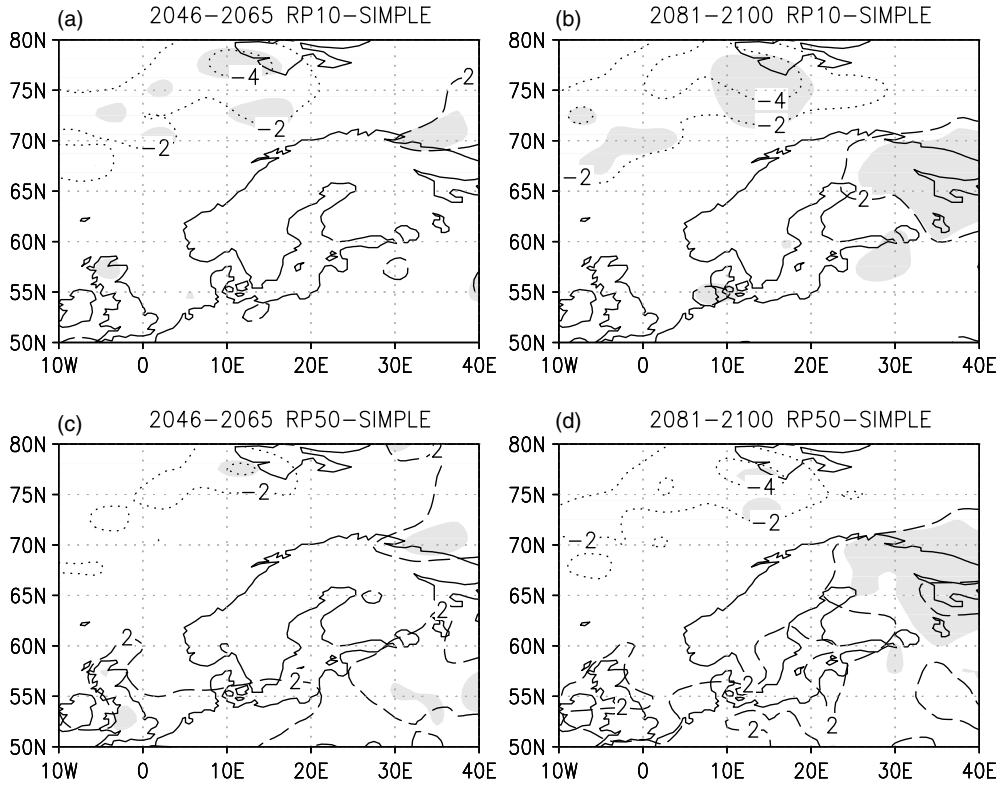

Figure 7. The percent changes in September–April extreme geostrophic wind speeds when using the simplified method with the time-dependent coefficients (Table III). For denotations see Fig. 5.

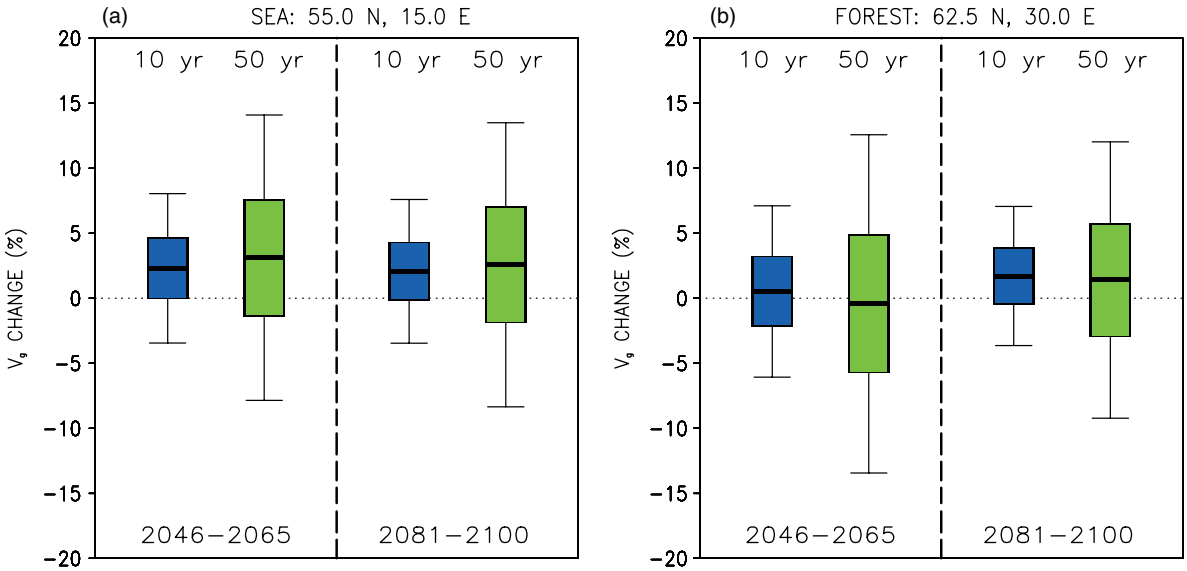

Figure 8. The probability distribution of the projected changes in the high-wind season (Sep–Apr) extreme geostrophic wind speeds from 1971–2000 to 2046–2065 and from 1971–2000 to 2081–2100 under the A1B : A2 : B1 scenario on the 10-year (dark shading) and 50-year (light shading) return levels. The bars represent the 25–75 and whiskers the 5–95 probability interval. Horizontal line inside the bars depicts the multimodel median. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

4. Discussion

In this work, the main focus was on analysing the projected changes of the mean and extreme geostrophic winds (V_g) in Northern Europe as well as the uncertainties and sensitivity in the analyses. The analyses of the extreme V_g on the different return periods were done using the GEV theory (Coles, 2001, Castillo *et al*., 2004) and the block maxima approach with the program R utilizing the extRemes software toolkit package (Katz *et al*., 2005). As the responses in extreme wind speeds were approximately homogeneously proportional to the average maximum wind speeds changes in September–April, a simplified method for estimating wind extremes was

Copyright © 2011 Royal Meteorological Society

developed. This method used the annual maxima statistics at two grid points (55°N, 15°E and 62.5°N, 30°E). Applying the simplified approach, the unrealistic return level estimates found at some grid points in the study area in some of the GCMs could be filtered.

The projected changes with the mean and extreme geostrophic winds became clearly significant in the latter 20-year period 2081–2100. For the mean wind the changes and their significance covered larger areas than the changes in the extreme wind speeds did (Figure 3(a),(b) *vs* Figure 5). However, in all the calculations the resulting pattern bared several similarities: there was an increase in southern and eastern Northern Europe and a decrease over the Norwegian Sea. The A1B and A2 forcing showed stronger increases at the end of the century than the B1 forcing.

By comparing the GCM-based average high-wind season maximum V_{gx} to the corresponding extreme geostrophic wind speeds on the various return periods V_{gx}(RP), we found that for instance the 10 and 50 year return period extreme wind speeds are, respectively, $13\% \pm 2\%$ and $22\% \pm 5\%$ stronger than the average September–April maxima. This finding was true for the three time periods and two grid point locations studied. The extreme wind speeds will increase on average by 2–4% in the southern and eastern parts of Northern Europe, whereas a decrease of 2–6% dominates over the Norwegian Sea. Nikulin *et al*. (2011), who evaluated the future projections of 20-year return level estimates of gust winds by performing the downscaling of six GCMs, found similarly that the increase in winds is dominant in a zone stretching from northern parts of France over the Baltic Sea towards northeast. In their work the decrease over the Norwegian Sea was somewhat less robust and stretched over Norway.

5. Conclusions and final remarks

The two methods used for analyzing the extreme wind speeds in this work gave qualitatively similar results. The results are more robust for the short return periods. Therefore, we assume that only a small part of the uncertainty in the estimated return level estimates originated from the extreme value analysis itself and that most of the uncertainty related to the extreme wind speed analyses resulted from the spread among the different GCMs. Similarly, Kharin *et al*. (2007), who studied the changes in temperature and precipitation extremes of a GCM ensemble, concluded that model differences generally dominate the uncertainty. Additionally, as the block maxima approach is somewhat sensitive to the sample size, a better way could be to employ the peak over threshold (POT) and bootstrapping methods (Naess and Clause, 2000) instead. Wehner (2010) who focused on the sources of uncertainty in the extreme temperature analyses found, however, that the return level estimates obtained by employing averaged empirical and parametric methods compare reasonably well even then when the cumulative distribution function is problematic to fit to the sample.

When we examine the individual months in Finland and the Baltic Sea, we can see that wind speeds are projected to increase from October to February. Despite the fact that the extreme winds are not expected to increase greatly in Northern Europe in the future, the wind-induced risks to forests could still be expected to increase under the changing climate. The influence of an increase in wind speed in forests in Finland is accompanied and strengthened by the reduction of the frozen soil period during the windiest seasons (i.e. from late autumn to early spring) (Peltola *et al*., 1999, Gregow *et al*., 2011). Moreover, under the warming climate the forest growth is projected to increase (Kellomäki and Leinonen, 2005, Kellomäki *et al*., 2008), which increases the need to manage forests more often or with higher intensity (i.e., thinning, final fellings) (Peltola *et al*., 2010). In the future work, similar analyses as presented in this paper would be important to do by examining the wind speeds together with other climate parameters. For instance, the changes in the wind speeds and directions, snow loads and soil frost (Gregow *et al*., 2008, 2011) could be combined model-wise for assessing the uncertainties in the projected climatological risks to wind- and snow-induced damages in forests. Wind speeds and snow loads in unfrozen and frozen ground have different impacts on the exposure of trees to bending, breakage and uprooting depending also on the tree age, stand characteristics, tree type and the recent management practices (Peltola *et al*., 1999, Gregow *et al*., 2011).

Although the GCM simulate the extratropical cyclone tracks and even the storm developments rather well (Geng and Sugi, 2003, Leckebusch and Ulbrich, 2004, Lambert, 2004, Bengtsson *et al*., 2006, Leckebusch *et al*., 2008, Ulbrich *et al*., 2008, Fink *et al*., 2009), there are still many details that need further research, especially when the focus is on winds. Aspects such as the projected reduction of snow cover (Jylhä *et al*., 2008, Räisänen and Eklund, 2011), retreat of sea-ice (Gordon *et al*., 2000, Mueter and Litzow, 2008), uncertainties in the land surface parameterization (Fischer *et al*., 2010) and the boundary layer processes as well as the modes in the ocean circulation can have essential unknown feedbacks to the storm and wind climate. As many such parameters and feedbacks are currently not yet adequately described in the GCMs, there is clearly still such uncertainty, i.e. in the climate projections that cannot even be estimated. Therefore, the results of this work should be considered mainly suggestive.

Acknowledgements

We acknowledge the modelling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling (WGCM) for making available the WCRP CMIP3 multi-model dataset. Seppo Saku is thanked for technical assistance. This work was financially supported by the Academy of Finland (2008–2010) via funding of the project on Impacts of temporal and spatial variability of

Copyright © 2011 Royal Meteorological Society

critical weather events and forest management on the risk of wind- and snow-induced damage in forest stands, led by Dr Heli Peltola, University of Eastern Finland, School of Forest Sciences. Additional funding was provided by the Finnish Ministry of Agriculture via Finnish Research Programme on Adaptation to Climate Change. Dr Ari Venäläinen from the Finnish Meteorological Institute is thanked for the initial discussions. An anonymous referee is acknowledged for several useful suggestions that have improved the paper significantly.

Appendix: Monthly mean V_g changes at various grid points in Finland and over the Baltic Sea

In 2046–2065, over the southern Baltic Sea, the V_g (Month) increases during October–November and January–February by 3–6% (Table AI). In Finland, such an increase is only seen in November while other months are only showing 1–3% changes of either positive or negative sign. Decrease dominates in Finland in March and April.

For the period 2081–2100, the changes are qualitatively similar to those for 2046–2065 but larger in amplitude. However, a general decrease is now only seen in March (Table AII). In November, the increase is 5–10%, being largest in the Baltic Sea and smallest in northern Finland. In January, the corresponding increase is 4–7%. More intense winds are likewise expected in October in the future although the change is mostly 2–4%. The differences between individual months indicate that an increase in the wind speeds is most likely in September–February.

Table AI. The 9-GCM average percentage changes in monthly mean wind speeds at the grid points presented in Table I as a response to the A1B scenario for the period 2046–2065 relative to 1971–2000.

2046–2065	Month											
Location	1	2	3	4	5	6	7	8	9	10	11	12
Baltic Sea	3	5	0	−1	1	4	−3	−1	0	4	6	2
Helsinki	2	3	−2	−1	0	2	0	−1	3	2	6	1
Joensuu	0	2	−2	−1	−1	−1	−1	−2	3	2	5	1
Jyväskylä	2	2	−3	−1	−1	0	1	−1	3	2	5	1
Kauhava	2	2	−2	−1	−1	1	0	0	3	2	4	1
Kajaani	1	1	−2	−2	−1	0	−2	−1	1	2	3	0
Rovaniemi	2	1	−2	−2	−1	1	0	0	1	1	3	0
Sodankylä	1	1	−2	−3	0	0	−1	0	0	2	3	0

Table AII. The 9-GCM average percentage changes in monthly mean wind speeds at the grid points presented in Table I as a response to the A1B scenario for the period 2081–2100 to 1971–2000.

2081–2100	Month											
Location	1	2	3	4	5	6	7	8	9	10	11	12
Baltic Sea	6	6	1	3	1	1	0	0	0	4	10	5
Helsinki	6	2	−2	1	0	1	1	1	2	3	6	2
Joensuu	6	2	−3	0	−1	−1	0	−1	2	3	7	2
Jyväskylä	7	2	−3	1	0	0	0	0	1	4	5	1
Kauhava	6	2	−2	1	0	0	0	0	1	4	5	0
Kajaani	5	3	−3	0	−1	0	−3	−1	1	2	5	1
Rovaniemi	5	3	−3	0	0	1	−1	0	1	3	5	0
Sodankylä	4	2	−2	−2	−2	0	−1	1	0	3	5	2

References

Alexandersson H. 2005. Den stora januaristormen 2005. *Väder och Vatten* **1**: 11 (In Swedish).

Bärring L, Fortuniak K. 2009. Multi-indices analysis of southern Scandinavian storminess 1780–2005 and links to interdecadal variations in the NW Europe-North Sea region. *International Journal of Climatology* **29**: 373–384.

Bengtsson A, Nilsson C. 2007. Extreme value modeling of storm damage in Swedish forests. *Natural Hazards and Earth System Sciences* **7**: 3518–3543.

Bengtsson L, Hodges KI, Rockner E. 2006. Storm tracks and climate change. *Journal of Climate* **19**: 3518–3543.

Bengtsson L, Hodges KI, Keenlyside N. 2009. Will extratropical storms intensify in a warmer climate? *Journal of Climate* **22**: 2276–2301.

Broyden CG. 1970. The convergence of a class of double-rank minimization algorithms. *Journal of the Institute of Mathematics and Its Applications* **6**: . 76–90.

Brüdl M, Rickli C. 2002. The storm Lothar 1999 in Switzerland – an incident analysis. *Forest Snow Landscape Research* **77**: 207–216.

Castillo E, Hadi AS, Balakrishnan N, Srabia JM. 2004. *Extreme Value and Related Models with Applications in Engineering and Science*. Wiley: New York.

Chang EKM, Lee SY, Swanson KL. 2002. Storm track dynamics. *Journal of Climate* **15**: 2163–2183.

Coles S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag: London, 242 pp.

Dobbertin M. 2002. Influence of stand structure and site factors on wind damage comparing the storms Vivian and Lothar. *Forest Snow Landscape Research* **77**: 187–204.

Donohoe A, Battisti DS. 2009. Causes of reduced Atlantic storm activity in a CAM3 simulation of the last glacial maximum. *Journal of Climate* **22**: 4793–4808.

Fink AH, Brücher T, Ermert V, Krüger A, Pinto JG. 2009. The European storm Kyrill in January 2007: synoptic evolution, meteorological impacts and some considerations with respect to climate change. *Natural Hazards and Earth System Sciences* **9**: 405–423.

Fischer EM, Lawrence DM, Sanderson BM. 2010. Quantifying uncertainties in projections of extremes – a perturbed land surface parameter experiment. *Climate Dynamics*. DOI: 10.1007/s00382-010-0915-y.

Fletcher R. 1970. A new approach to variable metric algorithms. *Computer Journal* **13**: 317–322.

Geng Q, Sugi M. 2003. Possible change of extratropical cyclone activity due to enhanced greenhouse gases and sulphate aerosols – study with high resolution AGCM. *Journal of Climate* **16**: 2262–2274.

Goldfarb D. 1970. A family of variable metric updates derived by variational means. *Applied Mathematics and Computation* **24**: 23–26.

Gregow H, Venäläinen A, Peltola H, Kellomäki S, Schultz D. 2008. Temporal and spatial occurrence of strong winds and large snowfalls amounts in Finland during 1961–2000. *Silva Fennica* **42**(4): 515–534.

Gregow H, Peltola H, Laapas M, Saku S, Venäläinen A. 2011. Combined occurrence of wind, snow loading and soil frost with implications for risks to forestry in Finland under the current and changing climatic conditions. *Silva Fennica* **45**(1): 35–54.

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA. 2000. The simulation of SST, sea ice extent and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics* **16**: 147–166, DOI: 10.1007/s003820050010.

Copyright © 2011 Royal Meteorological Society

Hanna E, Cappelen J, Allan R, Jonsson T, Le Blancq F, Lillington T, Hickley K. 2008. New insights into North European and North Atlantic surface pressure variability, storminess, and related climatic change since 1830. *Journal of Climate* **21**(24): 6739–6766, DOI: 10.1175/2008JCLI2296.1.

Jylhä K, Fronzek S, Tuomenvirta H, Carter TR, Ruosteenoja K. 2008. Changes in frost, snow and Baltic Sea ice by the end of the twenty-first century based on climate model projections for Europe. *Climatic Change* **86**: 441–462.

Jönsson A-M, Bärring L. 2010. Future climate impact on spruce bark beetle life-cycle in relation to uncertainties in regional climate model data ensembles. *Tellus A* **63**(1): 158–173, DOI: 10.1111/j.1600-0870.2010.00479.x.

Jönsson AM, Appelberg G, Harding S, Bärring L. 2009. Spatio-temporal impact of climate change on the activity and voltinism of the spruce bark beetle, *Ips typographus*. *Global Change Biology* **15**: 486–499.

Katz RW, Bruch GS, Parlange MB. 2005. Statistics of extremes: modelling ecological disturbances. *Ecology* **86**: 1124–1134.

Kharin VV, Zwiers FW, Zhang X, Hegerl GC. 2007. Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate* **20**: 1419–1444.

Kellomäki S, Leinonen S. 2005. *Management of European Forests Under Changing Climatic conditions*, Faculty of Forestry Research Notes 163. University of Eastern Finland: Joensuu, 427.

Kellomäki S, Peltola H, Nuutinen T, Korhonen KT, Strandman H. 2008. Sensitivity of managed boreal forests in Finland to climate change, with implications for adaptive management. *Philosophical Transactions of the Royal Society B* **363**: 2339–2349.

Lambert S. 2004. Changes in winter cyclone frequencies and strengths in transient enhanced greenhouse warming simulations using two coupled climate models. *Atmosphere-Ocean* **42**(3): 173–181.

Lau NC. 1988. Variability of the observed mid-latitude storm tracks in relation to low-frequency changes in the circulation pattern. *Journal of Atmospheric Sciences* **45**: 2718–2743.

Leckebusch GC, Ulbrich U. 2004. On the relationship between cyclones and extreme windstorm events over Europe under climate change. *Global and Planetary Change* **44**: 181–193.

Leckebusch GC, Donat M, Ulbrich U, Pinto JG. 2008. Mid-latitude cyclones and storms in an ensemble of European AOGCMs under ACC. *Clivar Exchanges* **13**(3): 3–15.

Matulla C, Schöner W, Alexandersson H, von Storch H, Wang XL. 2007. European storminess: late nineteenth century to the present. *Climate Dynamics* **31**: 125–130.

Meehl GA, Covey C, Delworth T, Latif M, McAvaney B, Mitchell JFB, Stouffer RJ, Taylor KE. 2007. The WCRP CMIP3 multimodel dataset: a new era in climate change research. *Bulletin of American Society* **88**: 1383–1394.

Mueter FJ, Litzow MA. 2008. Sea ice retreat alters the biogeography of the Bering Sea continental shelf. *Ecological Applications* **18**: 309–320, DOI: 10.1890/07-0564.1.

Naess A, Clause PH. 2000. The peaks over threshold and bootstrapping for estimating long return period design values. In *8th ASCE Speciality Conference on Probabilistic Mechanics and Structural Reliability*, PMC2000-151.

Nakićenović N, Alcamo J, Davis G, de Vries B, Fenhann J, Gaffin S, Gregory K, Grübler A, Jung TY, Kram T, La Rovere EL, Michaelis L, Mori S, Morita T, Pepper W, Pitcher H, Price L, Riahi K, Roehrl A, Rogner H-H, Sankovski A, Schlesinger M, Shukla P, Smith S, Swart R, van Rooijen S, Victor N, Dadi Z. 2000. *IPCC Special Report on Emissions Scenarios*. Cambridge University Press: Cambridge and New York, 599 pp.

Nikulin G, Kjellström E, Hansson U, Strandberg G, Ullerstig A. 2011. Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations. *Tellus* **63A**: 41–55, DOI: 10.1111/j.1600-0870.2010.00466.x

Nykänen M-L, Peltola H, Quine C, Kellomäki S, Broadgate M. 1997. Factors affecting snow damage of trees with particular reference to European conditions. *Silva Fennica* **31**(2): 193–213.

Pellikka P, Järvenpää E. 2003. Forest stand characteristics and wind and snow induced forest damage in boreal forest. In *Proceedings of the International Conference on Wind Effects on Trees*, 16–18 September 2003, University of Karlsruhe: Germany.

Peltola H, Kellomäki S, Väisänen H. 1999. Model computations on the impact of Climatic Change on the windthrow risk of trees. *Climatic Change* **41**: 17–36.

Peltola H, Ikonen V-P, Gregow H, Strandman H, Kilpeläinen A, Venäläinen A, Kellomäki S. 2010. Impacts of climate change on the forest dynamics and timber production with implications on the regional risks of wind-induced damage to forests in Finland. *Forest Ecology and Management* **260**(5): 833–845.

Pinto JG, Fröhlich EL, Leckebusch GC, Ulbrich U. 2007. Changing European storm loss potentials under modified climate conditions according to ensemble simulations of the ECHAM5/MPI-OM1 GCM. *Natural Hazards and Earth System Sciences* **7**: 165–175.

Räisänen J. 2007. How reliable are the climate models? *Tellus A* **59**(1): 2–29, DOI: 10.1111/j.1600-0870.2006.00211.x

Räisänen J, Eklund J. 2011. 21st century changes in snow climate in Northern Europe as simulated by regional climate models in the ENSEMBLES project: a high- resolution view from ENSEMBLES regional climate models. *Climate Dynamics*. DOI: 10.1007/s00382-011-1076-3.

Rõõm R. 1998. Acoustic filtering in nonhydrostatic pressure coordinate dynamics: a variational approach. *Journal of the Atmospheric Sciences* **55**(4): 654–668.

Schulz J-P. 2008. Revision of the turbulent gust diagnostics in the COSMO model. *COSMO Newsletter* **8**: 17–22. http://www.cosmo-model.org/. (Accessed 19 May 2011).

Schönenberger W. 2002. Windthrow research after the 1990 storm Vivian in Switzerland: objectives, study sites, and projects. *Forest Snow Landscape Research* **77**: 9–16.

Shanno DF. 1970. Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation* **24**: 647–656.

Ulbrich U, Fink AH, Klawa M, Pinto JG. 2001. Three extreme storms over Europe in December 1999. *Weather* **56**: 70–80.

Ulbrich U, Pinto JG, Kupfer H, Leckebusch GC, Spangehl T, Reyers M. 2008. Changing northern hemisphere storm tracks in and ensemble of IPCC climate change simulations. *Journal of Climate* **21**: 1669–1679.

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf J-F, Morcrette J-J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* **612**: 2961–3012, DOI: 10.1256/qj.04.176

Valinger E, Fridman J. 1997. Modeling probability of snow and wind damage to Scotts pine stands using tree characteristics. *Forest Ecology Management* **97**: 215–222.

Wehner MF. 2010. Sources of uncertainty in the extreme value statistics. *Extremes* **13**: 205–217, DOI 10.1007/s10687-010-0105-7.

Wehner MF, Smith RL, Bala G, Duffy P. 2010. The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model. *Climate Dynamics* **34**: 241–247.

Wernli H, Durren S, Lininger MA, Zelly M. 2002. Dynamical aspects of the life cycle of the winter storm 'Lothar' (24–26 December 1999). *Quarterly Journal of the Royal Meteorological Society* **128**: 405–430.

Zilitinkevich S, Baklanov A, Rosti J, Smedman A-S, Lykosov V, Calanca P. 2002. Diagnostic and prognostic equations for the depth of the stably stratified Ekman boundary layer. *Quarterly Journal of the Royal Meteorological Society* **128**: 25–46, DOI: 10.1256/00359000260498770.

Copyright © 2011 Royal Meteorological Society

Paper II

© Venäläinen, A., Korhonen, N., Hyvärinen, O., Koutsias, N., Xystrakis, F., Urbieta, I. R, and Moreno, J. M. 2014. CC Attribution 3.0 License.

Reprinted, with permission from

Natural Hazards and Earth System Sciences, 14, 1477–1490,

doi: 10.5194/nhess-14-1477-2014

Nat. Hazards Earth Syst. Sci., 14, 1477–1490, 2014
www.nat-hazards-earth-syst-sci.net/14/1477/2014/
doi:10.5194/nhess-14-1477-2014
© Author(s) 2014. CC Attribution 3.0 License.

Natural Hazards and Earth System Sciences
Open Access

Temporal variations and change in forest fire danger in Europe for 1960–2012

A. Venäläinen[1], N. Korhonen[1], O. Hyvärinen[1], N. Koutsias[2], F. Xystrakis[2], I. R. Urbieta[3], and J. M. Moreno[3]

[1]Finnish Meteorological Institute, Climate Change Research Unit, P.O. Box 503, 00101 Helsinki, Finland
[2]Department of Environmental and Natural Resources Management, University of Ioannina, G. Seferi 2, 30100 Agrinio, Greece
[3]Department of Environmental Sciences, University of Castilla-La Mancha, 45071 Toledo, Spain

Correspondence to: A. Venäläinen (ari.venalainen@fmi.fi)

Received: 10 September 2013 – Published in Nat. Hazards Earth Syst. Sci. Discuss.: 6 November 2013
Revised: 7 April 2014 – Accepted: 23 April 2014 – Published: 11 June 2014

Abstract. Understanding how fire weather danger indices changed in the past and how such changes affected forest fire activity is important in a changing climate. We used the Canadian Fire Weather Index (FWI), calculated from two reanalysis data sets, ERA-40 and ERA Interim, to examine the temporal variation of forest fire danger in Europe in 1960–2012. Additionally, we used national forest fire statistics from Greece, Spain and Finland to examine the relationship between fire danger and fires. There is no obvious trend in fire danger for the time period covered by ERA-40 (1960–1999), whereas for the period 1980–2012 covered by ERA Interim, the mean FWI shows an increasing trend for southern and eastern Europe which is significant at the 99 % confidence level. The cross correlations calculated at the national level in Greece, Spain and Finland between total area burned and mean FWI of the current season is of the order of 0.6, demonstrating the extent to which the current fire-season weather can explain forest fires. To summarize, fire risk is multifaceted, and while climate is a major determinant, other factors can contribute to it, either positively or negatively.

1 Introduction

During the last century, global average surface temperatures in the world have increased by about 0.78 °C (IPCC, 2013). However, this change in surface temperature has not been uniform across global regions. In Europe the decadal mean surface temperature has risen by about 1.3 °C from pre-industrial times until 2002–2011 (Haylock et al., 2008; EEA, 2012). During the last 50 years, warming has been most pronounced in the Iberian Peninsula, central and north-eastern Europe, and in mountainous regions; for the Iberian Peninsula, this warming has been most evident in summer (Haylock et al., 2008; EEA, 2012). In addition, warm extremes have become more frequent and cold extremes more rare. As an example, Della-Marta et al. (2007) showed that, in western Europe, the average length of summer heat waves is now twice as long as in 1880, and is accompanied by a tripling in the occurrence of hot days. Precipitation changes are not uniform across Europe; while increases in this have been observed in the north (Scandinavia and the Baltic states) a decrease is apparent in some southern areas, like in the Iberian Peninsula and especially in north-western Spain and northern Portugal (EEA, 2012). Warming trends in the form of increasing daily air temperature extremes dominated central and western Europe, while in the same region there are indications of an insignificant increase in dry spells (Moberg and Jones, 2005). Additionally, Lloyd-Hughes and Saunders (2002) have drawn attention to drying tendencies over central Europe which are stronger for the winter period.

These observed changes in climate have the potential to affect fire danger and, ultimately, forest fire activity. In this context, it is interesting to determine whether fire-weather danger has changed in Europe during the last decades. Fire danger indices are commonly used to assess fire potential, and warnings are regularly issued by fire agencies based on them. Fire danger indices combine several relevant climatic/weather variables into suitable format that, for example, forest fires services use to organize their response to

forecasted risks (Andrews et al., 2003; Fujioka et al., 2008; Camia et al., 2010). Fire danger indices allow for comparisons between fire-season severity across years and regions. The Canadian Fire Weather Index, FWI, based on the Canadian Forest Fire Danger Rating System (Van Wagner, 1987), is one of the most widely used indices, in use in countries across North and South America, Europe and Asia (De Groot et al., 2006).

Temporal and spatial variation of fire danger has been studied, for example, by Camia et al. (2008) in Europe using the ERA Interim data set (Dee at al., 2011) and the Seasonal Severity Rating index based on FWI. According to their analyses, from 1981 to 2010 and onwards, there have been statistically highly significant increases in fire danger in southeastern Europe and the southern-most regions of the Iberian Peninsula. Mäkelä et al. (2012) used the fire danger index operational in Finland to estimate the long-term 1908–2011 temporal variation of fire danger in Finland. They found out that the year-to-year variation in fire danger was large, but no significant trend could be detected. Wastl et al. (2012) examined the long-term trends in meteorological forest fire danger in the Alps during the period 1951–2010 using several fire danger rating indices, and found significant increases in the western and southern Alps. The increase was quite small in the northern Alps, and no clear signal could be observed in the inner Alpine valleys. Bedia et al. (2012) tested different model reanalysis data sets in the calculation of fire danger for the Iberian Peninsula and found that, for most parts of Spain, the summer season fire danger has increased. They also noted that the results based on the ERA Interim data set (Dee et al., 2011) were more robust than the results based on the NCEP data set (Kistler et al., 2001).

The relationships between fire danger indices and forest fires in Europe have been explored in a number of studies. Koutsias et al. (2013) examined trends in air temperature and their relationship to forest fires in Greece during the years 1894–2010 and found that both increased through time, particularly during the last 40 years. The annual number of fires and area burnt was strongly correlated with the maximum temperatures and summer heat waves. An evaluation of FWI performed recently using fire observations of a 15-year period for Greece confirmed that the index is capable of predicting fire occurrence (Karali et al., 2014). Additionally, Dimitrakapoulos et al. (2011a) found that there was a positive correlation between the annual drought and fire occurrence in Greece during the years 1961–1997. Similar analysis relating downscaling methods and future projections of the various FWI in the Iberian Peninsula and Greece was performed by Bedia et al. (2013). Positive correlations between FWI values and fires have been found in other parts of the world, such as in Canada (Gillet et al., 2004).

High fire danger caused by dry, hot and windy weather increases the occurrence of fires. However, most ignitions of fires are caused by human activities. For example, according to Ganteume et al. (2013) only about 7 % of the fires (whose cause was known) in northern Europe, 0.5 % in central Europe and 5 % in southern Europe were caused by natural sources. Similarly, a review by Bowman et al. (2009) showed that the variation in the occurrence of fires cannot be predicted by climate forcing alone, but that other aspects such as human behaviour and the effectiveness of fire detection and suppression systems must also be taken into account. To summarize, weather conditions make the occurrence of fires possible and humans in most cases ignite the fire.

Though weather and climate are not the sole forcing mechanisms of forest fires, they determine the conditions for fires to occur and spread, once an ignition occurs, and this way remain as the key factors for explaining the spatial and temporal variability of fires at all scales, including the whole globe. Therefore, in this study the main objective is to investigate whether the recent changes in climate have had a discernible impact on weather-related fire danger in Europe during the past five decades. To answer this question we have examined the temporal variation of the danger as expressed by FWI due to the variation in meteorological conditions over different European regions during 1960–2012 using ERA-40 (Uppala et al., 2005) and ERA Interim (Dee et al., 2011) data sets. Additionally, we also have demonstrated the significance of fire-weather indices in forest fires in three test sites (Greece, Spain and Finland) by calculating the cross correlation between fire statistics and mean FWI values.

2 Material and methods

2.1 European level

Europe was divided into four regions (Fig. 1) and only the grid boxes containing land were studied. The area denoted as southern Europe includes part of northern Africa for this study. The southern region represents Mediterranean warm climate, the eastern region continental Eurasian climate, the western area more humid Atlantic climate, and the northern region cool Fennoscandian climate. In addition to regional mean values, the FWI trend was analysed for each grid square.

We used ERA-40 (Uppala et al., 2005) for the years 1960–1999 and ERA Interim (Dee et al., 2011) data sets for the years 1980–2012 to calculate the Canadian FWI for each year. ERA-40 and ERA Interim are atmospheric reanalysed data sets that are created by the European Centre for Medium-Range Weather Forecasts' (ECMWF) modelling system. These data sets are based on all available and appropriate observational data, such as weather stations, radio soundings, ship measurements, satellite measurements etc. The spatial resolution used in this study was $2.5° \times 2.5°$ and $1.5° \times 1.5°$ for ERA-40 and for ERA Interim respectively. The calculation of FWI was done for the whole year but the main parameters of interest were the March–September mean value of FWI, and the number of days when FWI was

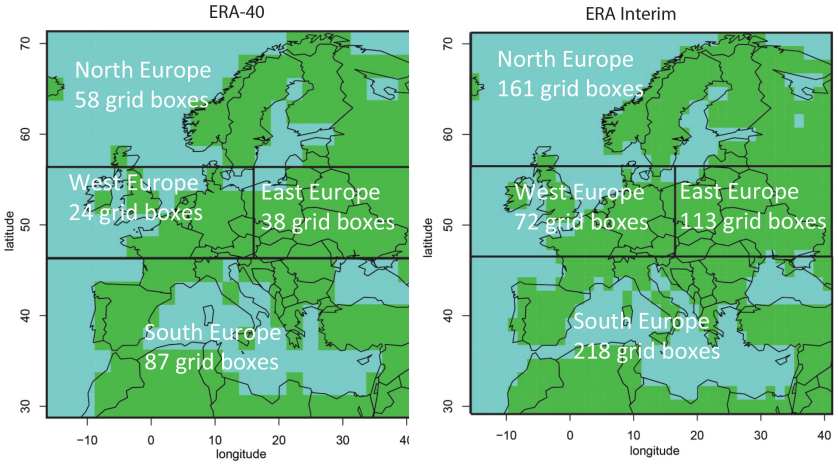

Figure 1. The four European regions analysed for ERA-40 (left panel) and ERA Interim (right panel) data sets.

larger than 20 or larger than 45 in March–September. The March–September period constitutes the main fire season in Europe. In northern Europe the season usually starts in May at the earliest after snow has melted, whereas in Southern Europe the season is centred in the summer months, although in some areas out-of-summer fires are also common. Moriondo et al. (2006) used an FWI threshold value of 45 for the Mediterranean region in their study on the impact of climate change on fire danger. Lehtonen et al. (2014) selected limits based on the study of Tanskanen et al. (2005) and the limits were FWI > 32 (extreme risk), 17–32 (high risk), 16–31 (medium risk) and < 8 low risk. The threshold of 20 used in this study can be regarded as applicable for the cool climate regions. In southern Europe, FWI values larger than 45 are relatively common, whereas in the rest of Europe they occur only very occasionally. This is why the analyses using the threshold of 45 were done only for the southernmost area and for the rest of the areas analyses were done using a threshold of 20. The results for the exceedance of threshold were presented as probabilities in percent (%):

$$\text{Probability} = 100 \times [\text{no. of cases}/(\text{no. grid boxes} \times \text{no. of days})]. \quad (1)$$

FWI has six components describing the moisture content of the surface layers, and the systems also predict various aspects of the fire behaviour. The input parameters required for deriving FWI are the midday temperature, relative humidity and wind speed, together with the precipitation sum of the previous 24 h. FWI is a dimensionless quantity indicating the fire danger. The FWI system was originally developed empirically for Canadian boreal conditions; however, the FWI indices have been proved to be good indicators of fire conditions in many other parts of the world, and have been proposed as the basis for a global early warning system for fires (De Groot et al., 2006).

In this study, the 12:00 UTC values were used. As the region used in the study is roughly from 5° W to 40° E, the 12:00 UTC value means four different solar times, which creates some inaccuracy in the calculations. However, as we are now interested in long-term temporal changes in FWI the exact numerical values are not that critical from that point as long as the method remains the same throughout the whole calculation period.

The calculation of FWI and trend of FWI was performed with the R package "fume" created by the Santander Meteorology Group (http://www.meteo.unican.es), which was used by Bedia et al. (2012). The trend was analysed using the Mann–Kendall test (Mann, 1945; Kendall, 1975) and the gradient of the trend line was calculated using Sen's slope estimate (Sen, 1968). The classical Mann–Kendall trend test evaluates the null hypothesis, H_0, that a time series is random (independent and identically distributed) against the alternative hypothesis, H_1, that the series exhibits a monotonic trend. However, the temporal autocorrelation can cause the classical test to reject the null hypothesis even if it is true. In its modified version, a correction factor is applied to the original variance formulation, accounting for the effective sample size in the presence of temporal autocorrelation (Bedia et al., 2012). This test is shown to be robust in the presence of serially correlated time series data (Hamed and Rao, 1998). Sen's slope estimate is the median of slopes calculated from all possible data value pairs. This estimate is more robust against outliers than, for example, least-squares regression. Trends were calculated separately for ERA-40 and ERA Interim data sets. In addition, to allow analyses covering the period 1960–2012, we estimated the ERA-40/ERA Interim coefficient from their common period (1980–1999) and extended the respective FWI time series on the basis of the estimated coefficient for the missing years. This resulted in two FWI time series, one corresponding to ERA-40 and one to ERA Interim for the period 1960–2012.

2.2 National level

Fire statistics at the national scalewere available for Greece for the period 1977–2010, for Spain for the period 1969–1999 and for Finland for the period 1960–2012. The Greek data comprise a subset from a national wildfire time series data (Koutsias et al., 2013) originally obtained from the National Statistical Service of Greece (NSSG), the Hellenic Forest Service (HFS), the Hellenic Fire Brigade (HFB) and Kailidis and Karanikola (2004). The Spanish data covered the 1961–2010 period and were obtained from the national forest fire statistics (EGIF, General Statistics of Wildfires) of the Spanish Ministry of Agriculture and Environment. The Finnish fire statistics have been published by the Finnish Forest Research Institute (2010).

For the national level calculations, in Spain, Greece and Finland we selected the FWI values to be analysed for the mid-summer months from June to September both from ERA-40 and ERA Interim data sets using only the grid cells located within peninsular Spain, Greece and Finland respectively. To identify possible abrupt shifts in the mean values of FWI in the time series, indicating distinct time periods, we applied the F statistic and the generalized fluctuation tests as described in Zeileis et al. (2003) and implemented in the R package "strucchange" (Zeileis et al., 2002). We applied both tests with a 0.1 bandwidth resulting in 5-year data window. To identify the optimal number of breakpoints, we adopted the method described in Zeileis et al. (2003). Within the defined segments, based on the breakpoints, we analysed the trend using the Mann–Kendall test, and the slope of the trend line was calculated using Sen's slope estimate.

Total burned area and number of fires were ln-transformed and the cross correlations with FWI were estimated applying the modified Pearson's correlation coefficient accounting for the autocorrelation of the time series using the approach followed by Meyn et al. (2010) by calculating the effective sample size that arises when a first-order correlation coefficient is considered. Additionally, cross correlations between FWI and the untransformed fire statistics were estimated by means of the non-parametric Spearman's coefficient also accounting for the autocorrelation in the time series. The cross-correlation analysis was performed using ± 3 lags (years) within the defined time segments in order to explore any bivariate lagged relationships between area burned and FWI.

3 Results

3.1 European level

3.1.1 Response of mean FWI

The temporal variation of mean FWI values since 1960 displays a relatively large year-to-year variation (Fig. 2). For the years 1980–1999 there is an overlap between ERA-40 and ERA Interim data sets, with an overall good agreement in the values of FWI from both. In southern Europe, the values based on ERA Interim are about the same as from ERA-40, whereas elsewhere FWI calculated from ERA-40 is systematically higher (Fig. 2).

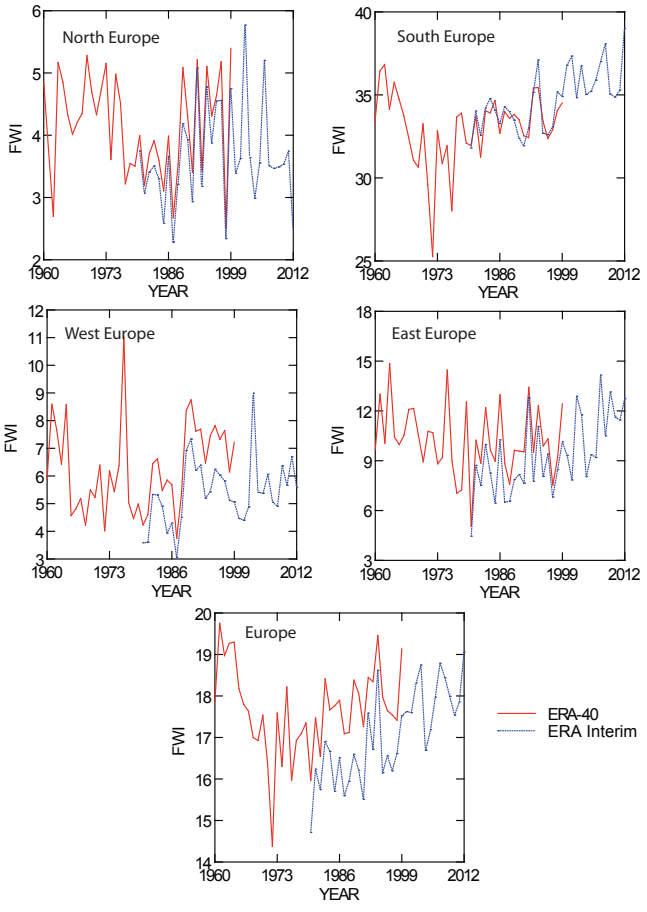

Figure 2. The year-to-year variation of March–September mean FWI from ERA-40 and ERA Interim data sets for four selected areas (see Fig. 1) and for the whole of Europe.

According to the trend analyses for ERA Interim-based FWI (Table 1), there was a significant upward trend at the 99 % level for southern and eastern Europe. The ERA Interim-based FWI trend for all of Europe also showed the same very high 99 % confidence level. However, ERA-40-based FWI did not exhibit any trend. The results show that during recent years, from approximately 1995 onwards, a tendency toward a higher FWI can be detected in the time series of Era Interim in all sub regions tested but the north. When we look at the trend for the whole period of 1960–2012 in the time series created by completing the ERA Interim backwards using ERA-40 data and the relationship defined using the common period of 1980–1999, the statistically significant rising trend at the 99 % level was found for the southern Europe. For eastern and western Europe the rising trend was statistically significant at 90 % level, and for

Table 1. Trend and gradient of the mean March–September FWI values as calculated for the regions of Europe in Fig. 1 using the modified Mann–Kendall test and Sen's method on ERA-40 and ERA Interim data sets. The following symbols indicate significance at level $\alpha = 0.01$ (***), $\alpha = 0.05$ (**) and $\alpha = 0.1$ (*).

Time series	First year	Last year	No. of years	Test Z	Signific.	Gradient
ERA-40 Europe	1960	1999	40	0.22		0.005
ERA-40 north	1960	1999	40	−0.69		−0.006
ERA-40 south	1960	1999	40	0.30		0.011
ERA-40 east	1960	1999	40	−1.13		−0.024
ERA-40 west	1960	1999	40	1.34		0.040
ERA Inter. Europe	1980	2012	33	4.54	***	0.087
ERA Inter. north	1980	2012	33	0.60		0.009
ERA Inter. south	1980	2012	33	5.58	***	0.116
ERA Inter. east	1980	2012	33	3.64	***	0.155
ERA Inter. west	1980	2012	33	1.65	*	0.041

northern Europe there was no trend. The spatial analysis of the trends (Fig. 3) showed that there was a very clear statistically significant rising trend in ERA Interim-based FWI in eastern Europe and also in the eastern Iberian Peninsula and much of France. A similar analysis for ERA-40-based FWI showed that there were statistically significant trends in a few cells only, with exception of the central Europe, in which there was a slight rising trend.

3.1.2 Response of FWI exceeding selected thresholds

The time series for the number of days when FWI exceeds 20 or 45 (Figs. 4 and 5) showed the same features as the time series for mean FWI based on Era Interim data (Fig. 2), with the number of days exceeding the selected threshold value becoming more frequent during the last 10 years or so. No trend was shown by ERA-40-based FWI, which was only available until 1999.

According to the trend analyses for the number of days where FWI > 20 (Table 2), ERA Interim-based FWI displays an upward trend significant at the 99 % level for southern Europe and at the 95 % level for eastern Europe. The trend shown for the whole of Europe also reaches the 99 % confidence level. For the number of days where FWI > 45 the trend is significant at the 99 % level for southern Europe. Again, there is no trend indicated by ERA-40-based FWI.

3.2 National level

3.2.1 Greece

The analysis of whether the mean FWI based on ERA-40 and ERA Interim data for Greece changes over time, based on the F statistic and the generalized fluctuation tests, did not result in any statistically significant breakpoint. However, there is a distinct period of low FWI values (Fig. 6). In order to objectively assess the time window of the period with distinctly low values of FWI, the breakpoint analysis (forcing for two breaks) resulted in the years 1970 and 1976. Trend analysis using the Mann–Kendall test and Sen's method did not reveal any significant positive or negative trend.

The cross-correlation analysis between the ERA-40 and ERA Interim-based FWI and the untransformed and transformed (ln) total area burned at the national scale in Greece within the period of 1977–2010 resulted in significant lag 0 coefficients (Fig. 7). Their values, as shown in Fig. 7, are 0.70 and 0.67 for ERA Interim and 0.69 and 064 for ERA-40 for the untransformed and the ln-transformed values respectively. The correlation coefficients for the number of fires are considerably smaller than those of total burned area (see the Supplement).

3.2.2 Spain

The analysis of whether the mean FWI based on ERA-40 and ERA Interim data for Spain changes over time resulted in three breakpoints, in the years 1968, 1977 and 1999, therefore defining four distinct periods – i.e. 1960–1968, 1969–1977, 1978–1999 and 2000–2011 – for both data sets (Fig. 8). Here, it should be pointed out that the 1977 breakpoint could be an outcome of the combination between the extremely low FWI values observed during 1971 and 1972 (Fig. 8) and the short bandwidth that was selected for the analysis (5 years). Within the three distinct periods, none of the data series was found to have a significant trend using the Mann–Kendall test and Sen's method. However, in the period defined when the last three sub-periods were merged (1969–2011), FWI of both data sets, ERA-40 and ERA Interim, as modified for completing the missing years, showed significant positive trends at significance level of 0.001 similar to the positive trends observed of FWI time series data of ERA-40 and ERA Interim south. In the case of the ERA-40-based FWI, a positive trend is observed when the original values that cover the period 1969–2001 ($p = 0.001$), while in the case of ERA Interim-based FWI, a positive trend is observed when the original values are used covering the period 1980–2012 ($p = 0.01$).

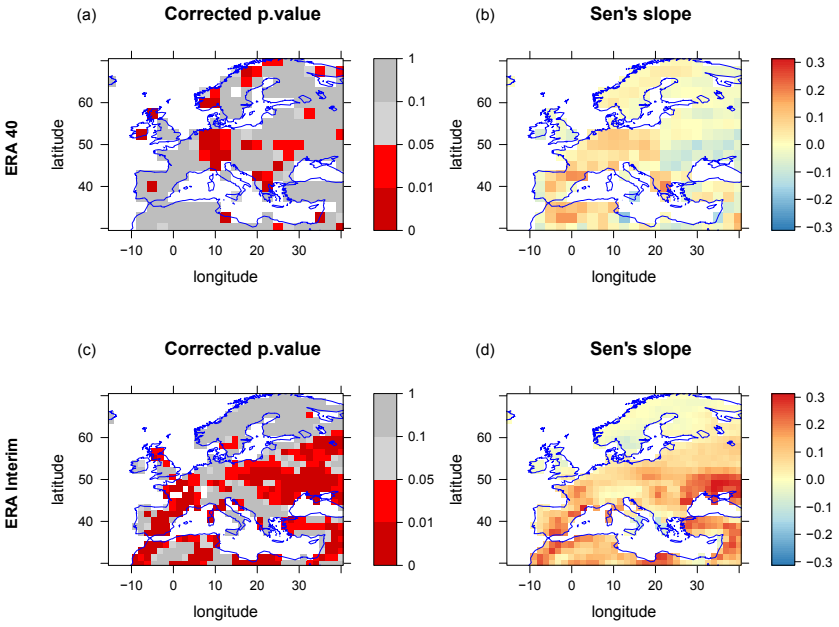

Figure 3. Trend of March–September mean FWI calculated using the ERA-40 data set for 1960–1999 **(b)** and the ERA Interim data set for 1980–2012 **(d)**. The statistical significance (α) is shown at levels 0.01, 0.05 and 0.1 for ERA-40 **(a)** and ERA Interim **(c)**.

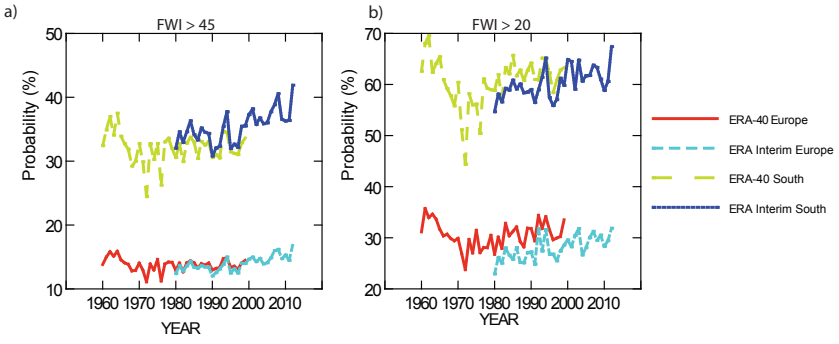

Figure 4. The year-to-year variation of March–September FWI values above 20 and above 45 from ERA-40 and ERA Interim data sets calculated for southern Europe and for the whole of Europe. The value is expressed as a probability (Eq. 1).

For the cross-correlation analysis we defined the period from 1969 to 1999 (i.e. not considering the breakpoint of 1978). The cross-correlation graphs between FWI of ERA-40 and ERA Interim south and the untransformed and transformed (ln) total area burned at the national scale in Spain (Fig. 9) indicate significant correlations at lag 0. The correlation coefficients, as shown in Fig. 9, are 0.50 and 0.62 for ERA Interim and 0.59 and 0.67 for ERA-40 for the untransformed and the ln-transformed burned area values respectively. The correlation coefficients for the number of fires are smaller than those of total burned area, though not as much as in the case of Greece (see the Supplement).

3.2.3 Finland

No trend in FWI was found for Finland, only large year-to-year variation (Fig. 10). The correlation between the burned area and FWI was roughly as high as in the case of the two Mediterranean countries studied above, i.e. around 0.6 (Fig. 11). Specifically, the correlation coefficients, as shown in Fig. 11, are 0.63 and 0.59 for ERA Interim and 0.61 and 0.57 for ERA-40 for the untransformed and the ln-transformed burned area values respectively. The correlation coefficients for the number of fires are smaller than those of total burned area, though not as much as in the case of Greece (see the Supplement).

Table 2. Trend and gradient of the March–September FWI values higher than 20 or higher than 45 as calculated for the different areas of Europe (Fig. 1) using the modified Mann–Kendall test and Sen's method and ERA-40 and ERA Interim data sets. The following symbols indicate significance at level $\alpha = 0.01$ (***), $\alpha = 0.05$ (**) and $\alpha = 0.1$.

Time series	First year	Last year	No. of years	Test Z	Signific.	Gradient
ERA-40 Europe > 20	1960	1999	40	0.72		0.004
ERA-40 south > 20	1960	1999	40	0.60		0.025
ERA-40 north > 20	1960	1999	40	−1.11		−0.037
ERA-40 west > 20	1960	1999	40	1.16		0.094
ERA-40 east > 20	1960	1999	40	−1.22		−0.095
ERA-40 Europe > 45	1960	1999	40	−0.90		−0.011
ERA-40 south > 45	1960	1999	40	−0.16		−0.006
ERA Inter. Europe > 20	1980	2012	33	3.98	***	0.161
ERA Inter. south > 20	1980	2012	33	3.08	***	0.166
ERA Inter. north > 20	1980	2012	33	0.64		0.045
ERA Inter. west > 20	1980	2012	33	1.53		0.081
ERA Inter. east > 20	1980	2012	33	3.33	**	0.391
ERA Inter. Europe > 45	1980	2012	33	4.11	***	0.078
ERA Inter. south > 45	1980	2012	33	4.01	***	0.145

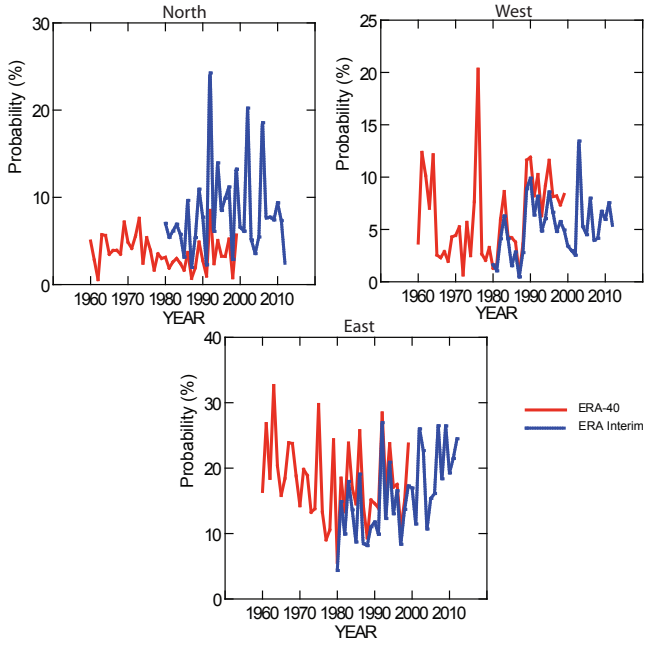

Figure 5. The year-to-year variation of March–September FWI values above 20 from ERA-40 and ERA Interim data sets calculated for north, west and east. Values are expressed as a probability (Eq. 1).

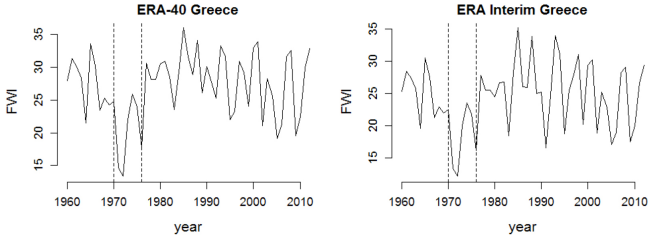

Figure 6. FWI time-series data based on ERA-40 and ERA Interim data for Greece. Vertical dotted lines correspond to forced breakpoints.

4 Discussion

The main purpose of this study was to detect the climate change signal in fire danger and to assess its importance for actual fire occurrence in two contrasting European regions, the Mediterranean and the boreal region, using various countries as examples. It is made apparent that the FWI trends observed in this study are compatible with the general trend towards climatic features characterized by warmer temperatures and changed precipitation patterns in Europe (Della-Marta et al., 2007; Haylock et al., 2008; EEA, 2012). The rise of fire danger during the last 40 years has occurred in southern and eastern Europe. In northern and western Europe temperature rise during summer season has not been that significant, and while precipitation has no decreasing trend, the net effect is that FWI did not exhibit a long-term trend either.

The temporal variations of FWI also depicts the long-term temporal variations of European climate; that is, early 1960s FWI had relatively high values in southern and southeastern Europe, followed by lower values in the late 1960s and early 1970s, and after that the change is characterized by a rising trend. An interesting detail is the area covering the northern side of the Black Sea, where FWI has a decreasing trend when calculated using ERA-40 and a notable increasing trend in ERA Interim. These variations demonstrate that although climate is warming, the long-term climate variations will remain, and also in the future, periods with lower FWI can be experienced even in areas characterized by an increasing trend of FWI.

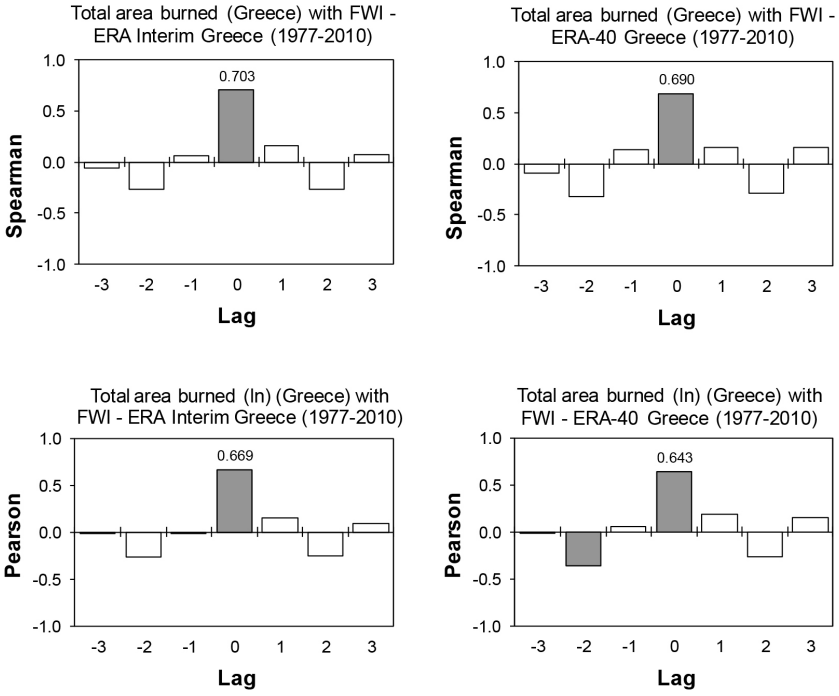

Figure 7. Cross-correlation graphs between total burned area (original and ln-transformed) at the national scale in Greece and FWI values estimated from ERA-40 and ERA Interim Greek data for the period 1977–2010 (grey columns indicate significant values at the 95 % confidence level).

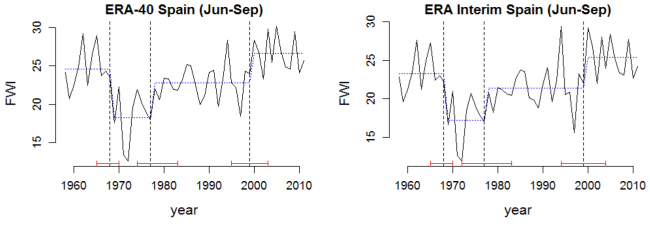

Figure 8. Breakpoints of FWI time-series data based on ERA-40 and ERA Interim for Spain. Vertical dotted lines correspond to breakpoints, while red horizontal lines correspond to the 95 % confidence intervals of the estimated breakpoints. Blue lines represent the mean values of FWI for each identified segment.

More specifically, and focusing on the three countries studied, it has been shown that, generally and with significant variability among regions and seasons, rainfall amount in the Greek peninsula shows a trend of decline after the 1980s, which is confirmed from various sources (Maheras et al., 2000; Pnevmatikos and Katsoulis, 2006; Feidas et al., 2007). This further supports the argumentation of Dimitrakopoulos et al. (2011c) regarding an increase in summer drought episodes. In Spain, temperatures have increased, although not homogeneously, across the country (Brunet et al., 2007; Pérez and Boscolo, 2010; Fernández-Montes et al., 2013; Acero et al., 2014), while precipitation changes, although less robust, also evidence a tendency towards a reduction, particularly during the last half of the century in the lower south-eastern half of the country (Pérez and Boscolo, 2010; Beguería et al. 2011). Although the annual mean temperature has risen in Finland by around 1 °C since 1900, the summer season mean temperature has not risen during the past 50 years (Tietäväinen et al., 2010). Similarly, there is no trend in summer precipitation. This lack of trend in either temperature or precipitation is reflected in temporal variation of FWI calculated for Finland; that is, there is no trend, only large year-to-year variation typical for northern Europe's highly variable climate.

The identification of breakpoints in the various series analysed helped in defining homogenous periods in the FWI time series that could be used to test trend-free cross correlations with fire statistics, rather than making a thorough analysis of the observed FWI variability. It is interesting to observe that, although the analysis did not identify any significant breakpoint, the pattern between the FWI series from Greece and Spain is very similar and characterized by a distinct period of low FWI values during the 1970s. In Spain, this period begins earlier and the differences in FWI values between this period and the time period just before are higher. This is the main reason for the identification of significant breakpoints in the Spanish data but not in the Greek data. However, respective breaks in the mean air temperature time series over various regions in Greece in the early 1970s and middle 1990s (Nastos et al., 2011; Kolokythas and Argiriou, 2013) and in

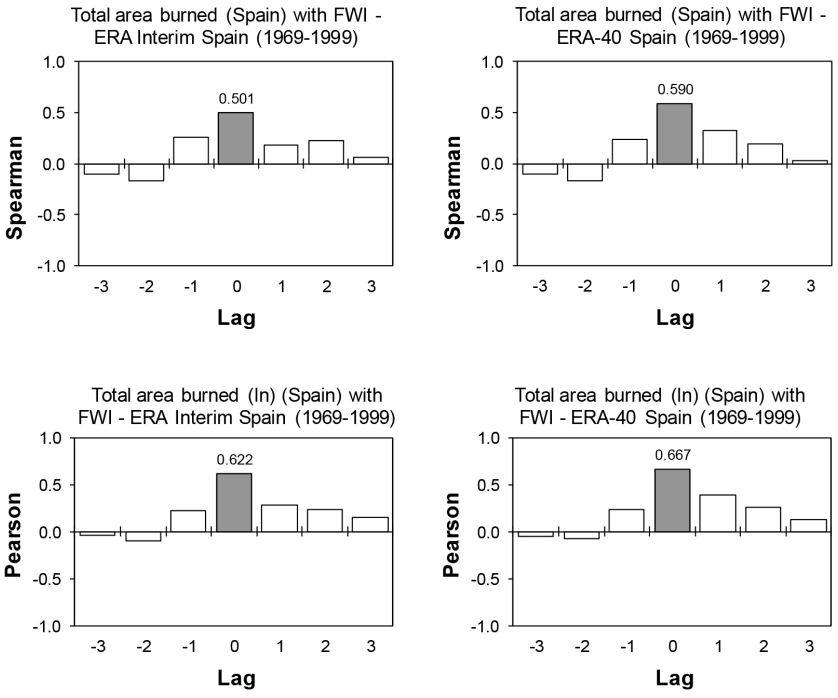

Figure 9. Cross-correlation graphs between total burned area (original and ln-transformed) at the national scale in Spain and FWI values estimated from ERA-40 and ERA Interim Spain data for the period 1969–1999 (grey columns indicate significant values at the 95 % confidence level).

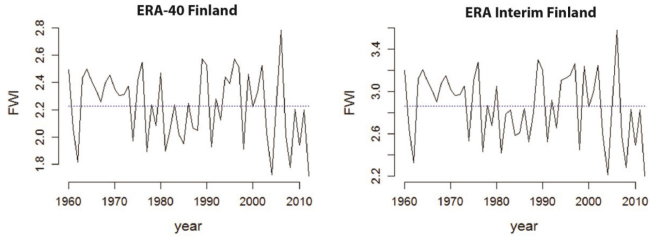

Figure 10. FWI time-series data based on ERA-40 and ERA Interim for Finland. Blue lines represent the mean values of FWI.

the precipitation time series between 1970 and 1980 (in the broader sense) (Pnevmatikos and Katsoulis, 2006; Feidas et al., 2007; Kalimeris et al., 2012) lead to the conclusion that an actual break in FWI series in the decade between 1970 and 1980 is also plausible. It should be noticed, however, that further testing is required in order to identify whether the breakpoints are artificial and due to exogenous variables or whether they actually indicate altering periods in the FWI values.

Correlations between fire statistics and FWI were apparent in all studied countries (Spain, Greece and Finland). As was shown here, the correlations between FWI and fire statistics were relatively high, and especially for Greece, they were similar to those estimated by Dimitrakopoulos et al. (2011b). This demonstrates, on the one hand, the causal effect of meteorological conditions on forest fires and, on the other hand, the importance of other factors such as vegetation/fuel, landscape and socioeconomics, including fire suppression, that are not related to climatic or meteorological conditions (Viedma et al., 2006, 2009; Carmo et al., 2011; Koutsias et al., 2013; Martínez-Fernández et al., 2013; Brotons et al., 2013). In all studied countries, Spain, Greece and Finland, the correlation coefficients of the number of fires are smaller than those of total burned area (see the Supplement). This can reflect in part the fact that fire statistics are more susceptible to being affected by small-size fires in particular. Area burned, however, is less subject to these spurious effects owing to the fact that large fires account for a high percentage of the total area burned, and statistics should be more robust in this regard (Moreno et al., 2011a; Pereira et al., 2011).

The profound and tight relation between weather and fire statistics has been demonstrated in numerous studies covering various parts and climate regions of the globe. Extreme fire weather is related to large fires in boreal ecosystems and sub-alpine forests (Beverly and Martell, 2005; Bessie and Johnson, 1995; Drobyshev et al., 2012). Similarly, in the Mediterranean region, the positive links between fire statistics and various weather components that indicate drought are apparent. Variables that most often correlate with burned area include simple weather parameters such as fire season precipitation (Vázquez and Moreno, 1993; Holden et al., 2007; Koutsias et al., 2013; Xystrakis and Koutsias,

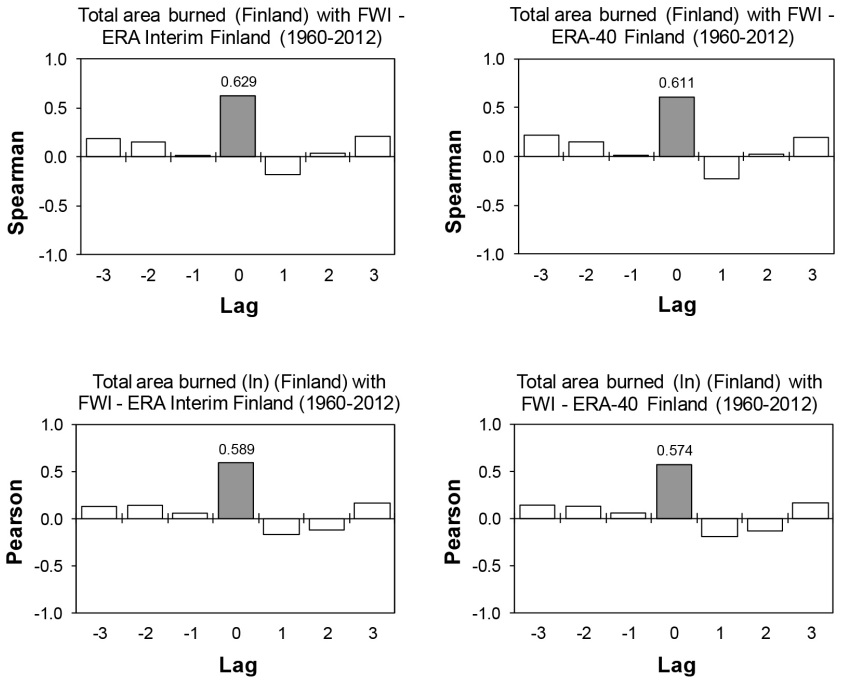

Figure 11. Cross-correlation graphs between total burned area (original and ln-transformed) at the national scale in Finland and FWI values estimated from ERA-40 and ERA Interim Finnish data for the period 1960–2012 (grey columns indicate significant values at the 95 % confidence level).

2013; Pausas, 2004), air temperature (Vázquez and Moreno, 1995; Piñol et al., 1998) or indices which, combining various weather parameters, quantify drought and fire danger (Carvalho et al., 2008; Camia and Amatulli, 2009). Although correlation does not necessarily imply causation, the links are strong and consistent, and the outputs of the present study also support this argumentation through the observed significant correlations. Such correlations between weather conditions and fire activity (including fire ignition and spread) are based on the control of weather in processes related to fuel moisture content and to the effect of wind in fire spread (Camia and Amatulli, 2009; Sullivan, 2009).

Large fires may account for more than 70 % of the total annual area burned (Ganteaume and Jappiot, 2013), and weather may be the dominant factor in determining fire spread (Moreira et al., 2011), with wind speed being identified as the main weather component of large forest fires in Greece (Dimitrakopoulos et al., 2011a) and California (Moritz et al., 2010); however, a large part of variation remains unexplained if other parameters are excluded. The lack of topographic, socioeconomic and landscape parameters may hamper correlations, and this could partially be the reason for the relatively low correlation coefficients between FWI values and fire statistics. Such additional factors were found to play a major role in southern France (Ganteaume and Jappiot, 2013), where area burned is related to high vegetation cover. In central Spain, landscape variability (discontinuity) was an important parameter for controlling fire size, even if the vast majority of the fire events took place under severe fire weather with FWI larger than 30 (Viedma et al., 2009). The influence of landscape in fire spread is also indirectly revealed through the analysis of selectivity patterns of fires belonging to different size classes (Moreira et al., 2011). Large fires in Sardinia tend to selectively burn shrub and grassland (Bajocco and Ricotta, 2008), and in central Spain large fires showed selectivity towards conifers and areas near settlements and roads (Moreno et al., 2011b). Socioeconomic variables like unemployment and touristic pressure could be directly linked to area burned (Koutsias et al., 2010; Ganteaume and Jappiot, 2013) or major social processes, as, for example, rural migration and urbanization followed by land abandonment could indirectly favour fuel conditions that could lead to large fire events as long as the fire is initiated (Koutsias et al., 2012). Similarly, humid-cool weather conditions could control fire initiation and spread even in the most fire prone ecosystems, indicating the dual role of weather in controlling fire size (Xystrakis et al., 2014). It can be argued that fire weather triggers or inhibits the landscape, topographic and socioeconomic variables in emerging as dominant factors of fire spread (Bradstock, 2010; Moreira et al., 2011).

5 Conclusions

In this study we examined the temporal variation of climate and fire weather danger in Europe using FWI. The results show an increase in fire danger in southern and eastern Europe after around 1970, and no clear signal elsewhere in Europe. This is consistent with observed patterns of climate change across the continent. Additionally, we found that fire danger indices and area burned were cross-correlated in Greece, Spain and Finland, although the correlations were not very high. Moreover, fires were only related to current-year climate variables. Though the meteorological conditions influencing on fire danger have changed in some areas, forest fires have not necessarily followed suit. For example, in the Mediterranean region, burned area has decreased, though FWI values show an increasing trend. This is consistent with the fact that weather and climate are major factors controlling fires, but not the only ones. The prediction of the occurrence of fires during the coming decades requires, in addition to climate research, comprehensive knowledge of socioeconomic aspects influencing fires and fire suppression. Furthermore, the foreseen ecological changes, like the changes in forest structures and dominant tree species, must be taken into account when the future fire conditions are estimated. The versatility of this problem emphasizes the importance of cross-disciplinary forest fire research.

The Supplement related to this article is available online at doi:10.5194/nhess-14-1477-2014-supplement.

Acknowledgements. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 243888 (FUME project). The authors also wish to express their cordial thanks to Ewan O'Connor for linguistic revision. ECMWF ERA-40 and ERA Interim data used in this study have been obtained from the ECMWF data server.

Edited by: P. Tarolli
Reviewed by: four anonymous referees

References

Acero, F., García, J., Gallego, M., Parey, S., and Dacunha-Castelle, D.: Trends in Summer Extreme Temperatures over the Iberian Peninsula using Non-Urban Station Data, J. Geophys. Res. Atmos., 119, 39–53, doi:10.1002/2013JD020590, 2014.

Andrews, P., Loftsgaarden, D., and Bradshaw, L.: Evaluation of fire danger indexes using logistic regression and percentile analysis, Int. J. Wildlad Fire, 12, 213–226, 2003.

Bajocco, S. and Ricotta, C.: Evidence of selective burning in Sardinia (Italy): which land-cover classes do wildfires prefer?, Landscape Ecol., 23, 241–248, doi:10.1007/s10980-007-9176-5, 2008.

Bedia, J., Herrera, S., Gutiérrez, J. M., Zavala, G., Urbieta, I. R., and Moreno, J. M.: Sensitivity of fire weather index to different reanalysis products in the Iberian Peninsula, Nat. Hazards Earth Syst. Sci., 12, 699–708, doi:10.5194/nhess-12-699-2012, 2012.

Bedia, J., Herrera, S., Martín, D. S., Koutsias, N., and Gutiérrez, J. M.: Robust projections of Fire Weather Index in the Mediterranean using statistical downscaling, Climatic Change, 120, 229–247, 10.1007/s10584-013-0787-3, 2013.

Beguería, S., Angulo-Martínez, M., Vicente-Serrano, S., López-Moreno, J., and El-Kenawy, A.: Assessing trends in extreme precipitation events intensity and magnitude using non-stationary peaks-over-threshold analysis: a case study in northeast Spain from 1930 to 2006, Int. J. Climatol., 31, 2102–2114, 2011.

Bessie, W. C. and Johnson, E. A.: The relative importance of fuels and weather on fire behavior in sub-Alpine forests, Ecology, 76, 747–762, doi:10.2307/1939341, 1995.

Beverly, J. L. and Martell, D. L.: Characterizing extreme fire and weather events in the Boreal Shield ecozone of Ontario, Agr. Forest Meteorol., 133, 5–16, doi:10.1016/j.agrformet.2005.07.015, 2005.

Bowman, D., Balch, K., Artaxo, P., Bond, W., Carlson, J., Cochrane, M., D'Antonio, C., DeFries, R., Doyle, J., Harrison, S., Johnston F., Keeley, J., Krawchuk, M., Kull, C., Marston, J., Moritz, M., Prentice, C., Roos, C., Scott, A., Swetnam, T., Der Werf, G., and Pyne, S.: Fire in the Earth System. Science, 324, 481–484, 2009.

Bradstock, R. A.: A biogeographic model of fire regimes in Australia: current and future implications, Global Ecol. Biogeogr., 19, 145–158, doi:10.1111/j.1466-8238.2009.00512.x, 2010.

Brotons, L., Aquilué, N., de Cáceres, M., Fortin, M., and Fall, A.: How Fire History, Fire Suppression Practices and Climate Change Affect Wildfire Regimes in Mediterranean Landscapes, PLoS ONE 8, e62392, 2013.

Brunet, M., Jones, P., Sigró, J., Saladié, O., Aguilar, E., Moberg, A., Della-Marta, P., Lister, D., Walther, A., and López, D.: Temporal and spatial temperature variability and change over Spain during 1850–2005, J. Geophys. Res. Atmos., 112, D12117, doi:10.1029/2006JD008249, 2007.

Camia, A. and Amatulli, G.: Weather factors and fire danger in the Mediterranean, in: Earth Observation of Wildland Fires in Mediterranean Ecosystems, edited by: Chuvieco, E., Springer Berlin Heidelberg, 71–82, 2009.

Camia, A., Amatulli, G., and San-Miguel-Ayanz, J. Past and future trends in forest fire danger in Europe, JRC Scientific and Technical Reports, EUR 23427 EN, 7 pp., Ispra, Italy, 2008.

Camia, A., Durrant Houston, T., and San-Miguel, J.: The European Fire Database: Development, Structure and Implementation, in: Proc. VI International Conference on Forest Fire Research, edited by: Viegas, D. X., Coimbra, 2010.

Carmo, M., Moreira, F., Casimiro, P., and Vaz, P.: Land use and topography influences on wildfire occurrence in northern Portugal, Landscape Urban Plan., 100, 169–176, doi:10.1016/j.landurbplan.2010.11.017, 2011.

Carvalho, A., Flannigan, M. D., Logan, K., Miranda, A. I., and Borrego, C.: Fire activity in Portugal and its relationship to weather and the Canadian Fire Weather Index System, Int. J. Wildland Fire, 17, 328–338, doi:10.1071/Wf07014, 2008.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Hólm, E., Isaksen, L., Kåll-

berg, P., Köhler, M., Matricardi, M., McNally, A., Monge-Sanz, B. M., Morcrette, J., Park, B., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137, 553–597, 2011.

De Groot, W., Goldammer. J., Keenan, T., Brady, M., Lynham., T., Justice, C., Csiszar, I., and O'Loughlin, K.: Developing a global early warning system for wildland fire, in: Proceedings of V International Conference on Forest Fire Research, edited by: Viega, D., 2006.

Della-Marta, P., Luterbacher, J., von Weissenfluh, E., Xoplaki, E., Brunet, M., and Wanner, H.: Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability, Clim. Dynam., 29, 251–275, doi:10.1007/s00382-007-0233-1, 2007.

Dimitrakopoulos, A., Gogi, C., Stamatelos, G., and Mitsopoulos, I.: Statistical analysis of the fire environment of large forest fires (> 1000 ha) in Greece, Polish J. Environ. Studies, 20, 327–332, 2011a.

Dimitrakopoulos, A. P., Bemmerzouk, A. M., and Mitsopoulos, I. D.: Evaluation of the Canadian fire weather index system in an eastern Mediterranean environment, Meteorol. Appl., 18, 83–93, doi:10.1002/met.214, 2011b.

Dimitrakopoulos, A. P., Vlahou, M., Anagnostopoulou, C. G., and Mitsopoulos, I. D.: Impact of drought on wildland fires in Greece: implications of climatic change?, Climatic Change, 109, 331–347, doi:10.1007/s10584-011-0026-8, 2011c.

Drobyshev, I., Niklasson, M., and Linderholm, H. W.: Forest fire activity in Sweden: Climatic controls and geographical patterns in 20th century, Agr. Forest Meteorol., 154–155, 174–186, doi:10.1016/j.agrformet.2011.11.002, 2012.

EEA: Climate change, impacts and vulnerability in Europe, An indicator-based report. EEA report 12/2012, Luxembourg: Office for Official Publications of the European Union, 2012 ISBN 978-92-9213-346-7, ISSN 1725-9177 doi:10.2800/66071, 2012.

Feidas, H., Noulopoulou, C., Makrogiannis, T., and Bora-Senta, E.: Trend analysis of precipitation time series in Greece and their relationship with circulation using surface and satellite data: 1955–2001, Theor. Appl. Climatol., 87, 155–177, 2007.

Fernández-Montes, S., Rodrigo, F., Seubert, S., and Sousa, P.: Spring and summer extreme temperatures in Iberia during last century in relation to circulation types, Atmos. Res., 127, 154–177, 2013.

Finnish Forest Research Institute: Finnish Statistical Yearbook of Forestry. Finnish Forest Research Institute, 472 pp., Helsinki, Finland, 2010 (in Finnish with summary and captions in English).

Fujioka, F., Gill, A., Viegas, D., and Wotton, B.: Fire Danger and Fire Behavior Modeling Systems in Australia, Europe, and North America, in: Developments in Environmental Sciences, edited by: Bytnerowicz, A., Arbaugh, M. J., Riebau, A. R., and Andersen, C., 471–497, Elsevier, Netherlands and United Kingdom, 2008.

Ganteaume, A. and Jappiot, M.: What causes large fires in Southern France, Forest Ecol. Manage., 294, 76–85, doi:10.1016/j.foreco.2012.06.055, 2013.

Ganteaume, A., Camia, A., Jappiot, M., San-Miguel-Ayanz, J., Long-Fournel, M., and Lampin, C.: A review of the main driving facors of forest fire ignition over Europe, Environ. Manage., 51, 651–662, 2013.

Gillet, N., Weaver, A., Zwiers, F., and Flannigan, M.: Detecting the effect of climate change on Canadian forest fires, Geophys. Res. Lett., 31, L18211, doi:10.1029/2004GL020876, 2004.

Hamed, K. H. and Rao, A. R.: A modified Mann Kendall trend test for autocorrelated data, J. Hydrol., 204, 182–196, 1998.

Haylock, M., Hofstra, N., Klein Tank, A., Klok, E., Jones, P., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res., 113, D20119, doi:10.1029/2008JD010201, 2008.

Holden, Z. A., Morgan, P., Crimmins, M. A., Steinhorst, R. K., and Smith, A. M. S.: Fire season precipitation variability influences fire extent and severity in a large southwestern wilderness area, United States, Geophys. Res. Lett., 34, 1–5, doi:10.1029/2007gl030804, 2007.

IPCC: Summary for Policymakers, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Kailidis, D. and Karanikola, P.: Forest fires 1990–2000, Giachoudi-Giapouli, Thessaloniki, 2004.

Kalimeris, A., Founda, D., Giannakopoulos, C., and Pierros, F.: Long-term precipitation variability in the Ionian Islands, Greece (Central Mediterranean): climatic signal analysis and future projections, Theor. Appl. Climatol., 109, 51–72, doi:10.1007/s00704-011-0550-5, 2012.

Karali, A., Hatzaki, M., Giannakopoulos, C., Roussos, A., Xanthopoulos, G., and Tenentes, V.: Sensitivity and evaluation of current fire risk and future projections due to climate change: the case study of Greece, Nat. Hazards Earth Syst. Sci., 14, 143–153, doi:10.5194/nhess-14-143-2014, 2014.

Kendall, M. G.: Rank correlation methods, Griffin, London, 1975.

Kistler, R., Kalnay, E., Collins,W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki,W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, B. Amer. Meteorol. Soc., 82, 247–268, 2001.

Kolokythas, K. V. and Argiriou, A. A.: Homogenization of Temperature Time Series of Western Greece, in: Advances in Meteorology, Climatology and Atmospheric Physics, edited by: Helmis, C. G. and Nastos, P. T., Springer Atmospheric Sciences, Springer Berlin Heidelberg, 535–540, 2013.

Koutsias, N., Martinez-Fernandez, J., and Allgower, B.: Do factors causing wildfires vary in space? Evidence from geographically weighted regression, GISci. Remote Sens., 47, 221–240, doi:10.2747/1548-1603.47.2.221, 2010.

Koutsias, N., Arianoutsou, M., Kallimanis, A. S., Mallinis, G., Halley, J. M., and Dimopoulos, P.: Where did the fires burn in Peloponnisos, Greece the summer of 2007? Evidence for a synergy of fuel and weather, Agr. Forest Meteorol., 156, 41–53, 2012.

Koutsias, N., Xanthopoulos, G., Founda, D., Xystrakis, F., Nioti, F., Pleniou, M., Mallinis, G., and Arianoutsou, A.: On the relationship between forest fires and weather conditions in Greece from long-term national observations (1894–2010), Int. J. Wildland Fire, 22, 493–507, 2013.

Lehtonen, I., Ruosteenoja, K., Venäläinen, A., and Gregow, H.: The projected 21st century forest fire risk in Finland under different greenhouse gas scenarios, Boreal Environ. Res., 19, 127–139, 2014.

Lloyd-Hughes, B. and Saunders, M. A.: A drought climatology for Europe, Int. J. Climatol., 22, 1571–1592, doi:10.1002/joc.846, 2002.

Maheras, P., Patrikas, I., Karacostas, T., and Anagnostopoulou, C.: Automatic classification of circulation types in Greece: methodology, description, frequency, variability and trend analysis, Theor. Appl. Climatol., 67, 205–223, doi:10.1007/s007040070010, 2000.

Mäkelä, H. M., Laapas, M., and Venäläinen, A.: Long-term temporal changes in the occurrence of a high forest fire danger in Finland, Nat. Hazards Earth Syst. Sci., 12, 2591–2601, doi:10.5194/nhess-12-2591-2012, 2012.

Mann, H. B.: Nonparametric tests against trend, Econometrica, Econometrica, 13, 245–259, 1945.

Martínez-Fernández, J., Chuvieco, E., and Koutsias, N.: Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression, Nat. Hazards Earth Syst. Sci., 13, 311–327, doi:10.5194/nhess-13-311-2013, 2013.

Meyn, A., Schmidtlein, S., Taylor, S. W., Girardin, M. P., Thonicke, K., and Cramer, W.: Spatial variation of trends in wildfire and summer drought in British Columbia, Canada, 1920–2000, Int. J. Wildland Fire, 19, 272–283, 2010.

Moberg, A. and Jones, P. D.: Trends in indices for extremes in daily temperature and precipitation in central and western Europe, 1901–99, Int. J. Climatol., 25, 1149–1171, doi:10.1002/joc.1163, 2005.

Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., Koutsias, N., Rigolot, E., Barbati, A., Corona, P., Vaz, P., Xanthopoulos, G., Mouillot, F., and Bilgili, E.: Landscape – wildfire interactions in southern Europe: Implications for landscape management, J. Environ. Manage., 92, 2389–2402, doi:10.1016/j.jenvman.2011.06.028, 2011.

Moreno, M., Malamud, B., and Chuvieco, E.: Wildfire frequency-area statistics in Spain, Proc. Environ. Sci., 7, 182–187, 2011a.

Moreno, J. M., Viedma, O., Zavala, G., and Luna, B.: Landscape variables influencing forest fires in central Spain, Int. J. Wildland Fire, 20, 678–689, 2011b.

Moriondo, M., Good, P., Durao, R., Bindi, M., Giannakopoulos, C., and Corte-Real, J.: Potential impact of climate change on fire danger in the Mediterranean area, Clim. Res., 31, 85–95, 2006.

Moritz, M. A., Moody, T. J., Krawchuk, M. A., Hughes, M., and Hall, A.: Spatial variation in extreme winds predicts large wildfire locations in chaparral ecosystems, Geophys. Res. Lett., 37, L04801, doi:10.1029/2009gl041735, 2010.

Nastos, P. T., Philandras, C. M., Founda, D., and Zerefos, C. S.: Air temperature trends related to changes in atmospheric circulation in the wider area of Greece, Int. J. Remote Sens., 32, 737–750, doi:10.1080/01431161.2010.517796, 2011.

Pausas, J. G.: Changes in fire and climate in the eastern Iberian Peninsula (Mediterranean Basin), Climatic Change, 63, 337–350, 2004.

Pereira, M. G., Malamud, B. D., Trigo, R. M., and Alves, P. I.: The history and characteristics of the 1980–2005 Portuguese rural fire database, Nat. Hazards Earth Syst. Sci., 11, 3343–3358, doi:10.5194/nhess-11-3343-2011, 2011.

Pérez, F. and Boscolo, R.: Clima en España: Pasado, Presente y Futuro. Informe de evaluación del cambio climático regional, Clivar España, available at: www.clivar.es/files/informe_clivar_final.pdf (last access: 6 June 2014), 2010.

Piñol, J., Terradas, J., and Lloret, F.: Climate warming, wildfire hazard, and wildfire occurrence in coastal eastern Spain, Climatic Change, 38, 345–357, 1998.

Pnevmatikos, J. D. and Katsoulis, B. D.: The changing rainfall regime in Greece and its impact on climatological means, Meteorol. Appl., 13, 331–345, doi:10.1017/S1350482706002350, 2006.

Sen, P.: Estimates of the regression coefficient based on Kendall's tau, J. Am. Stat. Assoc., 63, 1379–1389, 1968.

Sullivan, A. L.: Wildland surface fire spread modelling, 1990–2007, 2: Empirical and quasi-empirical models, Int. J. Wildland Fire, 18, 369–386, 2009.

Tanskanen, H., Venäläinen, A., Puttonen, P., and Granström, A.: Impact of stand structure on surface fire ignition potential in *Picea abies* and *Pinus sylvestris* forests in southern Finland, Can. J. For. Res,. 35, 410–420, 2005.

Tietäväinen, H., Tuomenvirta, H., and Venäläinen, A.: Annual and seasonal mean temperatures in Finland during the last 160 years based on gridded temperature data, Int. J. Climatol., 30, 2247–2256, 2010.

Uppala, S., Kallberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R., Andersson, E., Arpe, K., Balmaseda, M., Beljaars, A., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B., Isaksen, L., Janssen, P., Jenne, R., McNally, A., Mahfouf, J., Morcrette, J., Rayner, N., Saunders, R., Simon, P., Sterl, A., Trenberth, K., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Q. J. R. Meteorol. Soc., 131, 2961–3012, 2005.

Van Wagner, C.: Development and structure of the Canadian Forest Fire Weather IndexSystem. Canadian Forestry Service, Ottawa, Ontario, Forestry Technical Report, 35, 37 pp., 1987.

Vázquez, A. and Moreno, J. M.: Sensitivity of fire occurrence to meteorological variables in Mediterranean and Atlantic areas of Spain, Landscape Urban Plan., 24, 129–142, doi:10.1016/0169-2046(93)90091-Q, 1993.

Vázquez, A. and Moreno, J. M.: Patterns of fire occurrence across a climatic gradient and its relationship to meteorological variables in Spain, in: Global Change and Mediterranean-Type 10 Ecosystems, edited by: Moreno, J. M. and Oechel, W. C., Ecological Studies 117, Springer-Verlag, New York, 408–434, 1995.

Viedma, O., Moreno, J. M., and Rieiro, I.: Interactions between land use/land cover change, forest fires and landscape structure in Sierra de Gredos (Central Spain), Environ. Conservation, 33, 212–222, 2006.

Viedma, O., Angeler, D. G., and Moreno, J. M.: Landscape structural features control fire size in a Mediterranean forested area of central Spain, Int. J. Wildland Fire, 18, 575–583, doi:10.1071/Wf08030, 2009.

Wastl, C., Schunk, C., Leuchner, M., Pezzatti, G., and Menzel, A. Recent climate change: Long-term trends in meteorological for-

est fire danger in the Alps, Agr. Forest Meteorol., 162–163, 1–13, 2012.

Xystrakis, F. and Koutsias, N.: Differences of fire activity and their underlying factors among vegetation formations in Greece, iForest – Biogeosci. Forest., 6, 132–140, doi:10.3832/ifor0837-006, 2013.

Xystrakis, F., Kallimanis, A. S., Dimopoulos, P., Halley, J. M., and Koutsias, N.: Precipitation dominates fire occurrence in Greece (1900–2010): its dual role in fuel build-up and dryness, Nat. Hazards Earth Syst. Sci., 14, 21–32, doi:10.5194/nhess-14-21-2014, 2014.

Zeileis, A., Leisch, F., Hornik, K., and Kleiber, C.: strucchange: An R Package for Testing for Structural Change in Linear Regression Models, J. Stat. Softw., 7, 1–38, 2002.

Zeileis, A., Kleiber, C., Krämer, W., and Hornik, K.: Testing and dating of structural changes in practice, Comput. Statistics Data Analysis, 44, 109–123, 2003.

Paper III

© Korhonen, N., Venäläinen, A., Seppä, H., and Järvinen, H. 2014. CC Attribution 3.0 License.

Reprinted, with permission from
Climate of the Past, 10, 1489–1500,
doi:10.5194/cp-10-1489-2014

Clim. Past, 10, 1489–1500, 2014
www.clim-past.net/10/1489/2014/
doi:10.5194/cp-10-1489-2014
© Author(s) 2014. CC Attribution 3.0 License.

Climate of the Past
Open Access

Statistical downscaling of a climate simulation of the last glacial cycle: temperature and precipitation over Northern Europe

N. Korhonen[1,3], **A. Venäläinen**[1], **H. Seppä**[2], **and H. Järvinen**[3]

[1]Finnish Meteorological Institute, Climate Service Centre, Helsinki, Finland
[2]University of Helsinki, Department of Geosciences and Geography, Helsinki, Finland
[3]University of Helsinki, Department of Physics, Helsinki, Finland

Correspondence to: N. Korhonen (natalia.korhonen@fmi.fi)

Received: 27 May 2013 – Published in Clim. Past Discuss.: 21 June 2013
Revised: 17 June 2014 – Accepted: 29 June 2014 – Published: 7 August 2014

Abstract. Earth system models of intermediate complexity (EMICs) have proven to be able to simulate the large-scale features of glacial–interglacial climate evolution. For many climatic applications the spatial resolution of the EMICs' output is, however, too coarse, and downscaling methods are needed. In this study we introduce a way to use generalized additive models (GAMs) for downscaling the large-scale output of an EMIC in very different climatological conditions ranging from glacial periods to current relatively warm climates. GAMs are regression models in which a combination of explanatory variables is related to the response through a sum of spline functions. We calibrated the GAMs using observations of the recent past climate and the results of short time-slice simulations of glacial climate performed by the relatively high-resolution general circulation model CCSM (Community Climate System Model) and the regional climate model RCA3 (Rossby Centre regional Atmospheric climate model). As explanatory variables we used the output of a simulation by the CLIMBER-2 (CLIMate and BiosphERe model 2) EMIC of the last glacial cycle, coupled with the SICOPOLIS (SImulation COde for POLythermal Ice Sheets) ice sheet model, i.e. the large-scale temperature and precipitation data of CLIMBER-2, and the elevation, distance to ice sheet, slope direction and slope angle from SICOPOLIS. The fitted GAMs were able to explain more than 96 % of the temperature response with a correlation of > 0.98 and more than 59 % of the precipitation response with a correlation of > 0.72. The first comparison with two pollen-based reconstructions of temperature for Northern Europe showed that CLIMBER-2 data downscaled by GAMs corresponded better with the reconstructions than did the bilinearly interpolated CLIMBER-2 surface temperature.

1 Introduction

Climate risk assessments and bioclimatic studies on a millennial timescale require regional climate data. Present state-of-the-art comprehensive atmospheric general circulation models (GCMs), coupled with modules simulating the biosphere and sea ice, are major tools for the study of past, present and future climates. The resolution of such global models is usually 100–300 km (Flato et al., 2013). The relatively high resolution and complex calculations make GCMs suitable for modelling the climate on a 100–1000 yr timescale. For routinely simulating time periods covering full glacial cycles, the GCMs are computationally far too demanding. Nevertheless, several GCMs have been used for time-slice simulations of the distant past climate with prescribed ice sheets (Renssen et al., 2005; Otto-Bliesner et al., 2006; Brandefelt and Otto-Bliesner, 2009; Liu et al., 2009; Kjellström et al., 2010; Singarayer and Valdes, 2010; Strandberg et al., 2011; Braconnot et al., 2011, 2012). Smith and Gregory (2012) used an AOGCM (atmosphere–ocean general circulation model) in a transient simulation of the last glacial cycle; however, in their work the ice sheet was prescribed and not coupled with the climate component of the model.

Earth system models of intermediate complexity (EMIC; Claussen et al., 2002; Petoukhov et al., 2005) provide a practical approach for simulating climate in timescales covering entire glacial cycles ($\sim 100\,000$ yr). A reasonable computing

time is achieved by reducing the problem's size (i.e. number of grid points) and model complexity. Full earth system models employ eddy-resolving general circulation models for atmosphere and ocean and are coupled with detailed submodels for bio-geochemical processes and cycles. EMICs, in contrast, are coarsely discretized and an atmosphere model is typically dynamical–statistical where mean flow is explicit and eddies are parameterized. Ocean models can be multibasin but zonally averaged. Despite these simplifications, in an EMIC intercomparison of eight different models (Petoukhov et al., 2005), the equilibrium and transient responses to a doubling of the atmospheric CO_2 concentration were within the range of corresponding GCM simulations. EMICs have been extensively utilized in simulating past glaciations (Berger et al., 1999, 2003; Wang and Mysak, 2002; Calov et al., 2005a, b) and future climate on a timescale of 100 000 yr (Berger and Loutre, 2002; Archer and Ganopolski, 2005; Cochelin et al., 2006).

The EMIC simulations provide continuous global data; however, there are uncertainties related to the model simplifications and low spatial resolution, and these make the direct output of EMICs unsuitable for regional studies. In climate science, several downscaling methods have been used. A common method is to downscale the results of a global model dynamically with a regional model. Regional models usually have sophisticated atmosphere and biosphere modules at a resolution of ~ 50 km and are therefore computationally demanding. Hence downscaling of this kind is possible for short time-slice simulations of 50–100 yr, not for an entire 100 000 yr long simulation. A computationally less demanding method is statistical downscaling, whereby the statistical relationships between the observed small-scale variables (derived from observations) and larger-scale variables (e.g. from a global model) are derived using, for example, regression analysis. These derived statistical relations are then applied to downscale the large-scale variables of climate simulations to a smaller scale. Dynamical vs. statistical downscaling from hemispheric to regional scales involves a number of fundamental choices. It seems that the single most important factor affecting the quality of the downscaling is the simulation skill of the driving model (for a thorough discussion, see Racherla et al. (2012).

Vrac et al. (2007) introduced a statistical downscaling method for palaeoclimatological purposes, based on generalized additive models (GAMs; Wood, 2006). They fitted a GAM-type regression model by finding the statistical relationships between the observed recent past climate (1961–1990) and the low-resolution CLIMBER-2 (CLIMate and BiosphERe Model 2) EMIC (Petoukhov et al., 2000, 2005) model simulation of recent past climate. They then used this regression model to downscale the recent past and last glacial maximum climates simulated by CLIMBER-2 over Western Europe. All these GAMs were calibrated with the recent past climate, and were used relying on the assumption that the statistical relations between the large and small scales remain unchanged for the last glacial maximum and the recent past climate.

We investigated the usability of GAMs in downscaling the large-scale variables of the CLIMBER-2 EMIC model simulations by Ganopolski et al. (2010) over Europe in climatic conditions ranging from glacial to interglacial. For this purpose we fit a GAM not only to downscale the observed recent past climate but also to downscale relatively high-resolution simulated glacial climate. For a glacial period characterized by an extensive ice sheet over Fennoscandia, about 21 kyr before present (BP), we utilized data from a CCSM4 (Community Climate System Model) GCM simulation by Brady et al. (2013) and an RCA3 (Rossby Centre regional Atmospheric climate model) regional model simulation by Strandberg et al. (2011) forced by a CCSM3 simulation by Brandefelt and Otto-Bliesner (2009). For a glacial period characterized by a small ice sheet over Fennoscandia, about 44 kyr BP, we utilized data from an RCA3 simulation by Kjellström et al. (2010). RCA3 as a regional climate model has higher spatial resolution than the global model CCSM4 and in this study RCA3 is used because of its presumably more detailed depiction of climate, whereas CCSM4 is used because its output enables downscaling to be done for any location on earth. As stated by Vrac et al. (2007), the use of statistical downscaling in palaeoclimatology is based on the assumption that the fitted regression model remains unchanged over time. We evaluated the GAMs by presenting their skills and investigate thereby which explanatory variables produce the best fit. We also tested the GAMs for downscaling the temperature of the CLIMBER-2 simulation from 10 kyr BP up to the present. The bilinearly interpolated CLIMBER-2 surface temperature and the downscaled temperature were compared to two temperature reconstructions: Laihalampi, Finland (Heikkilä and Seppä, 2003), and Gilltjärnen, Sweden (Antonsson et al., 2006).

2 Downscaling with generalized additive models

Statistical regression models are calibrated by establishing statistical relationships between large-scale variables (called explanatory variables or predictors, $X_1, \ldots, X_p$) and a local variable (Y, called the response or predictand). Our large-scale data are the global CLIMBER-2 output (Ganopolski et al., 2010) covering the last glacial cycle, with climate conditions ranging from glacial to interglacial. The range of the large-scale variables, X, defines the calibration range. Hence, the calibration range of the regression model should cover glacial to interglacial climates. For local data representing interglacial climate we utilize observations of the recent past climate by the Climate Research Unit, CRU. For local data representing climate for certain time steps of the last glacial cycle we utilize simulations by a GCM (CCSM4; Gent et al., 2011) and a regional model (RCA3; Samuelsson et al., 2011). These two data sets were used separately. The

CCSM4 LGM (last glacial maximum) simulation (Brady et al., 2013) was used for the annual GAMs over the area we call Western Eurasia (36–70° N, 10°W–69° E). The RCA3 downscaling (Strandberg et al., 2011) of the CCSM3 LGM simulation (Brandefelt and Otto-Bliesner, 2009) was used in the monthly GAMs over Northern Europe (54–70° N, 3–35° E). We aim at finding the relationships between the high-resolution data and the large-scale data for the corresponding time periods. These relationships are then used with the large-scale data to obtain high-resolution data where they are otherwise not available.

2.1 Calibrating the statistical model

We used GAMs (Hastie and Tibshirani, 1990; Wood, 2006) in which the statistical expectation (E) of the predictand (Y) were modelled using the sum of univariate spline functions of the p predictors ($X_1, \ldots, X_p$), such that

$$E\left(Y|X_1 \ldots X_p\right) = \sum_{j=1}^{p} fj(Xj), \quad (1)$$

where the potentially non-linear spline functions f_j had a non-parametric form. The calibration of GAMs is data-driven, and the shape of the response is not forced to any parametric form. GAMs are called semiparametric models, as the probability distribution of the predictand should be known. Our predictand was either the mean temperature (temperature 2 m above the surface) or the mean total precipitation. The variables used for the calibration of the monthly GAMs were monthly means of each grid point. Respectively, the annual GAMs were calibrated using annual means of each grid point. As temperature data classically satisfy the normality assumption, the response variable mean temperature was expected to follow a normal distribution. According to Cheng and Qi (2002), the cumulative precipitation data can be modelled by a lognormal distribution, hence the total precipitation response variable was log-transformed and expected to follow a normal distribution. The statistical relationships were calibrated using the data for 1 month (or annual) at a time. For fitting the GAM, all the data were to be represented at the same spatial resolution. We used a resolution of $1.5° \times 0.75°$ over Northern Europe and Western Eurasia. These two regions are defined based on the available climate simulation data sets; the RCA3 simulations cover Northern Europe and the CCSM4 the area we call here Western Eurasia. As the size of the grid cells of our data varied with latitude, we associated a weight with each grid cell depending on the latitude of the data. The functions f_j in Eq. (1) were defined by cubic regression splines with a restricted maximum number of degrees of freedom under wiggliness penalization (Wood, 2006). This means that although the maximum number of the degrees of freedom per predictor was restricted here to 15, the penalization allowed reducing the effective number of degrees of freedom of each spline to what was really needed by the problem.

As predictors on the right-hand side of Eq. (1) we used both meteorological and geographic predictors. As meteorological predictors we tested CLIMBER-2 large-scale temperature, relative humidity, precipitation, and lapse-rate data. As geographic predictors we tested terrain properties (elevation, shortest distance to the ocean, distance to nearest water body, slope angle, direction of the slope with respect to north), latitude, longitude, and shortest distance to the ice sheet margin. We tested altogether about 50 combinations of predictors and found the present to be the best. All the predictors were physically reasonable. Ice sheets have a large impact on local climate and that is why one of the predictors tested was selected to be the shortest distance to the ice sheet margin. The predictor side's temperature, relative humidity, precipitation, lapse rate and, over ice-free areas, the terrain properties were extracted from the CLIMBER-2-SICOPOLIS (SImulation COde for POLythermal Ice Sheets) simulation data by Ganopolski et al. (2010) at 0, 44 and 21 kyr BP to represent the recent past, a stadial during MIS (marine isotope stage) 3 and LGM, respectively. Over ice-covered areas the terrain properties used as predictors were extracted from the data of the RCA3 and CCSM4 simulations. In earlier work GAMs have been developed using observed data only. However, if we want to have GAMs valid also for, for example, glacial periods, we need data depicting those conditions. The use of climate simulations as this work widens the data window and should hence lead to an outcome that covers various climate conditions better than the system that is developed using present data only.

The shortest distance to the nearest glacier margin used as predictor was computed from the RCA3 and CCSM4 ice sheet data. In ice-free areas the distance was expressed by positive values and in areas covered by ice the distance was expressed by negative values. In the calibration of the GAMs we implement the ice sheet extent and topography corresponding to the calibration data; e.g. when calibrating by RCA3 data we implement the RCA3 ice sheet extent and topography. Further, for predictions, the ice sheet extent and topography of SICOPOLIS is used. If the GAMs were calibrated by SICOPOLIS ice sheet data, we would have problems near the ice sheet margins as they differ in the SICOPOLIS simulation and RCA3 and CCSM4.

The final predictors were determined by minimization of the generalized cross-validation (GCV) score, which was computed as

$$\mathrm{GCV} = \frac{N\overline{(Y_\mathrm{m} - \overline{Y_\mathrm{m}})^2}}{(N - \gamma d)^2}, \quad (2)$$

where N stands for the number of observations, Y_m for the GAM-modelled surface air temperature or log-precipitation (log meaning natural logarithm), and d stands for the effective number of degrees of freedom. The GCV score is an estimate of the expected mean square predictor error inside the calibration predictor values' range. The GCV has the

same units as Y. Further, in computing the GCV, the effective degrees of freedom were also penalized by increasing their cost with a weight $\gamma = 1.4$ to prevent overfitting of the model (Kim and Gu, 2004). We allowed ourselves to combine several predictors, if that was found to improve the skill and was physically reasonable. We tested numerous combinations of predictors for fitting the GAMs in Eq. (1). Based on the GCV values, the best fits for downscaling the log-precipitation with GAMs were attained with the equation

$$\log(P)\,(P_{\mathrm{CLI}}, x, y, h, d) = s_1\,(P_{\mathrm{CLI}}) + s_2\,(x, y) + s_3\,(h) + s_4\,(d)\,, \quad (3)$$

in which P represents the downscaled precipitation, P_{CLI} is the precipitation of the large-scale data (CLIMBER-2), x and y are the longitude and latitude, respectively, h is the elevation, d is the direction of the steepest slope, and $s_1 \ldots s_4$ are spline functions.

For downscaling the annual temperature and the monthly temperature of May, June, July, August and September with GAMs, the best fits were attained with the equation

$$T\,(T_{\mathrm{CLI}}, x, y, h, i) = T_{\mathrm{CLI}} + s_5\,(x, y) + s_6\,(h, i)\,, \quad (4)$$

and for the monthly temperature of January, February, March, April, October, November and December with the equation

$$T\,(T_{\mathrm{CLI}}, x, y, h, d, s) = T_{\mathrm{CLI}} + s_7\,(x, y, T_{\mathrm{CLI}}) + s_8\,(h) + s_9\,(d, s)\,, \quad (5)$$

in which T is the downscaled temperature, T_{CLI} is the temperature of the large-scale data (CLIMBER-2), the slope, s, is the angle of the steepest slope, i is the shortest distance to the ice sheet margin (positive if ice-free, negative if covered with ice) and $s_5 \ldots s_9$ are spline functions.

2.2 Large-scale input data, CLIMBER-2-SICOPOLIS

The large-scale data to be downscaled were the output of a full last glacial cycle simulation (126 000 yr BP–present day) with the CLIMBER-2-SICOPOLIS model system by Ganopolski et al. (2010). CLIMBER-2 is an EMIC consisting of six earth system components: atmosphere, ocean, sea ice, land surface, terrestrial vegetation and ice sheets (Petoukhov et al., 2000). The model has a relatively low spatial resolution: for the atmospheric module the latitudinal resolution is 10° and the longitudinal resolution is roughly 51°. The CLIMBER-2 model successfully describes the seasonal variability of a large set of characteristics of the climate system (Petoukhov et al. 2000), and simulates the climate response to changes in different types of forcing and boundary conditions within the range of corresponding GCM simulations (Ganopolski et al., 2001; Petoukhov et al., 2005). For simulating glacial climates, the CLIMBER-2 model has been coupled with the high-resolution three-dimensional thermomechanical ice sheet model SICOPOLIS (Greve, 1997). The resolution of the ice sheet model is $1.5° \times 0.75°$ and the model domain extends in the Northern Hemisphere from 21 to 85.5° N. The climate and ice sheet components are coupled bidirectionally using a physically based energy and mass balance interface (SEMI) as described in detail by Calov et al. (2005a). The simulations by Ganopolski et al. (2010) were forced by variations in the earth's orbital parameters calculated following Berger (1978), and atmospheric greenhouse gas concentrations derived from the Vostok ice core. The evolution of the radiative forcing of atmospheric dust was parameterized to alter proportionally with the simulated global ice volume as by Schneider et al. (2006).

2.3 Regional-scale input data

2.3.1 Recent past climate

For the recent past climate predictands, we used the CRU high-resolution climate data, version 2.1. (Mitchell and Jones, 2005). For land areas the original resolution of the monthly mean surface temperature and total precipitation data was $0.5° \times 0.5°$. In practise this CRU data is the same as the often-used, reanalysed ERA-40 (ECMWF 40 Years Reanalysis) data set (Uppala et al., 2005). However, the use of ERA 40 or the other reanalysed data sets such as NCEP (National Centers for Environmental Prediction; Kistler et al., 2001) could be considered in the possible future research. To cover areas where CRU high-resolution data were not available, i.e. sea areas, we used the Jones et al. (1999) $5° \times 5°$ temperature data. Similarly, where high-resolution precipitation data were not available, we used the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP; Xie and Arkin, 1997) derived from the years 1979–2000, with the original data resolution of $2.5° \times 2.5°$.

2.3.2 Glacial climate over Northern Europe

For the glacial climate predictands over Northern Europe we used the 50 yr of monthly mean temperature and precipitation values of two simulations with the regional climate model RCA3 (Kjellström et al., 2005, Samuelsson et al., 2011) forced by simulations with the GCM CCSM3 (Collins et al., 2006; Kiehl et al., 2006). The first data set was that of an RCA3 simulation by Kjellström et al. (2010), representing a stadial within the MIS 3 around 44,000 yr BP, a cold period with a relatively small Fennoscandian ice sheet as illustrated in Fig. 1a. The second data set was that of an RCA3 simulation by Strandberg et al. (2011), forced with a CCSM3 simulation by Brandefelt and Otto-Bliesner (2009), representing the LGM around 21 000 yr BP, with an extensive ice sheet covering large parts of northern Eurasia, as illustrated in Fig. 1b.

Different phases of the Palaeoclimate Modelling Intercomparison Project (PMIP) provide detailed information on the relative performance of climate models in past climate

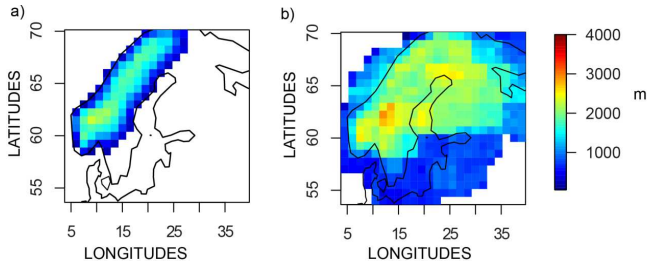

Figure 1. Fennoscandian ice sheet extent and elevation (m) **(a)** as in the RCA3 MIS 3 stadial simulation by Kjellström et al. (2010) and **(b)** as extracted from the ICE-5G data by Peltier (2004) for the RCA3 LGM simulation by Strandberg et al. (2011) . Data have been bilinearly interpolated to a 1.5° × 0.75° resolution.

conditions. In particular, the CCSM3 model compares generally well in the PMIP3 (Kageyama et al., 2006) and it is within the ensemble spanned by participating models; for instance, regarding the simulated temperature of the warmest and coldest month in the latitude–longitude sectorial averages of interest. We thus conclude that, for the present study, the choice of CCSM3 is equally well justified as the selection of any other model in the PMIP ensemble.

2.3.3 Glacial climate over Western Eurasia

The glacial climate predictands over Western Eurasia were produced using annual mean temperature and precipitation data from two simulations with GCM CCSM4: the last 20 yr of the CCSM4 LGM simulation (Brady et al., 2013) and the period AD 1961–1990 of the CCSM4 recent past simulation (Gent et al., 2011). First we computed the differences between the CCSM4 simulated LGM and recent past climates and then we added these changes to the CRU's annual temperature and precipitation data sets (Sect. 2.3.1). For temperature we applied the absolute change (°C) and for precipitation the relative change (%).

2.4 Evaluation of the GAMs

For evaluating the fitted GAMs, we computed several statistical quantities as follows:

(i) the percentage of explained deviance:

$$\%\mathrm{ED} = 100 * \frac{\overline{(Y_\mathrm{m} - \overline{Y_0})^2}}{(Y_0 - \overline{Y_0})^2}, \quad (6)$$

where Y_0 stands for the predictand data (CRU/RCA3/CCSM4) and Y_m for the GAM-modelled surface air temperature or log-precipitation;

(ii) the spatial correlation between the CRU/RCA3/CCSM4 and the GAM-modelled surface air temperature or log-precipitation:

$$\mathrm{Cor} = \frac{\sum_{i=1}^{n} (Y_\mathrm{m} - \overline{Y_\mathrm{m}})(Y_0 - \overline{Y_0})}{\sqrt{\sum_{i=1}^{n} (Y_\mathrm{m} - \overline{Y_\mathrm{m}})^2 \sum_{i=1}^{n} (Y_0 - \overline{Y_0})^2}}. \quad (7)$$

For error estimates we computed the root-mean-square error,

$$\mathrm{RMSE} = \sqrt{n^{-1} \sum_{i=1}^{n} (Y_0 - Y_\mathrm{m})^2}, \quad (8)$$

which is sensitive to the maximum errors, and the mean absolute difference,

$$\mathrm{MAD} = n^{-1} \sum_{i=1}^{n} |Y_0 - Y_\mathrm{m}|, \quad (9)$$

which is less sensitive to the maximum errors.

Further, in Supplement 1, we tested the monthly GAMs in predicting other months' temperatures or precipitation over Northern Europe. In Supplement 2, the annual GAMs for Western Eurasia (calibrated only by recent past and LGM data) were used for predicting annual mean temperature and precipitation for 44 kyr BP, and the results were compared to the output of the RCA3 regional climate model simulation by Kjellström et al. (2010).

2.5 Comparison with temperature reconstructions

The calibrated GAMs were applied to downscale the Holocene climate (10 kyr BP–present) of the simulation by Ganopolski et al. (2010). The resulting annual mean temperature was compared with two pollen-based reconstructions of temperature for Laihalampi, Finland (Heikkilä and Seppä, 2003), and Gilltjärnen, Sweden (Antonsson et al., 2006). The RMSE and MAD were computed for the bilinearly interpolated CLIMBER-2 surface temperature and reconstructions as well as for the downscaled temperature and reconstructions.

3 Results

3.1 Downscaling of precipitation

3.1.1 Evaluation of GAMs for precipitation

Equation (3) resulted in the skills presented in Table 1. The precipitation GAMs showed good correlation (> 0.72) with the fitted data. The MAD (RMSE) of the monthly precipitation GAMs varied between 7 mm month^{-1} (12 mm month^{-1}) in February and 12 mm month^{-1} (21 mm month^{-1}) in October. The GAMs were able to explain the spatial variance of the total precipitation by 59–85 %.

Figure 2 depicts the splines of the GAM downscaling January log-precipitation over Northern Europe. Figure 2a

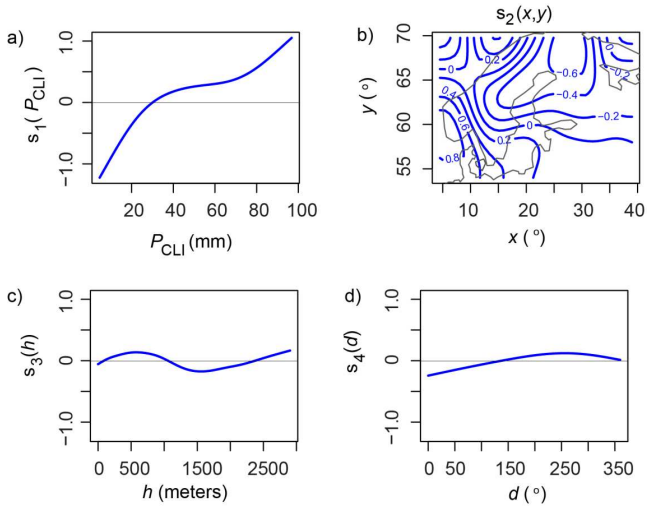

Figure 2. Components of the GAM model in Eq. (3) fitted to downscale the CLIMBER-2 January log-precipitation in the recent past, MIS 3 stadial and LGM climates. The blue curves show the estimated effects of **(a)** CLIMBER-2 total precipitation, **(b)** longitude and latitude, **(c)** elevation, and **(d)** direction of the steepest slope.

Table 1. Skills of the monthly total log-precipitation GAMs for the Northern European area, and of the annual total log-precipitation (∗) GAM for the Western Eurasian area. Values in parentheses represent total precipitation. The GCV score is defined in Eq. (2). The percentage of explained deviance (%ED), the spatial correlation (Cor), the root-mean-square error (RMSE), and mean absolute difference (MAD) are defined in Eqs. (6–10), respectively.

Month	GCV	%ED (%)	Cor	RMSE (mm month^{-1})	MAD (mm month^{-1})
January	0.09	84	0.91 (0.88)	14	8
February	0.09	82	0.90 (0.87)	12	7
March	0.08	79	0.89 (0.84)	13	7
April	0.08	67	0.82 (0.75)	13	8
May	0.06	63	0.80 (0.75)	13	8
June	0.06	63	0.80 (0.74)	14	10
July	0.06	67	0.82 (0.78)	15	11
August	0.06	59	0.77 (0.72)	17	12
September	0.07	63	0.79 (0.77)	20	12
October	0.08	72	0.85 (0.78)	21	12
November	0.10	78	0.88 (0.83)	18	11
December	0.10	81	0.90 (0.85)	18	10
Annual*	0.05	85	0.92 (0.84)	13	8

demonstrates the effect of the raw CLIMBER-2 January precipitation: the precipitation increased with CLIMBER-2's precipitation, which indicates that the spline function reduces the precipitation in months of low precipitation simulated with CLIMBER, and enhances the precipitation in months of high CLIMBER-simulated precipitation. This further leads to the conclusion that the spline function increases the amplitude of the precipitation annual cycle, which is a well justified assumption because of CLIMBER's coarse resolution and very small precipitation variability. Figure 2b shows the effect of the longitude and latitude: the spline increased precipitation on the west coast of Norway, as well as in Denmark and southern Sweden and decreased it over central and northern Finland. The effect of elevation on precipitation depended on the season. In the annual GAM and March–September GAMs, the precipitation increased with elevation (not shown). However, for cooler months, October–January, the elevation effect was weaker, and for elevations between 1000 and 2000 m the precipitation even decreased with elevation (Fig. 2c). To speculate, this may have a meteorological explanation: at relatively low elevations (< 750 m) topography enhances precipitation, between 750 and 2000 m elevations on glaciers the top climate is characterized by smaller precipitation amounts and, again, when we go above 2500 m (mountain tops) precipitation increases as a function of elevation. The effect of the direction of the slope is shown in Fig. 2d: southerly and westerly slopes (~ 180 to $\sim 270°$) tended to increase precipitation, whereas northerly and easterly slopes (~ 0 to $\sim 90°$) decreased it.

Table S1.1 in Supplement 1 shows the skills of the monthly GAMs in predicting other months' precipitation over Northern Europe; the correlations and errors are in the same range as for the fitting data, suggesting that these GAMs do make reasonable predictions with other than just calibration data.

3.1.2 Comparison of downscaled precipitation with observation

Figure 3 demonstrates the raw CLIMBER-2 January precipitation (top row in Fig. 3), the CRU (first panel of the second row in Fig. 3) and RCA3 (second and third panels of the second row in Fig. 3) regional-scale precipitation, the GAM downscaled January precipitation (third row in Fig. 3), and the difference between the CRU/RCA3 data and the GAM downscaled precipitation (bottom row in Fig. 3). By comparing rows three and two in Fig. 3, we see that the GAM reproduced well the spatial distribution of precipitation over Northern Europe, the highest precipitation rates occurring on the western slopes of the Scandinavian mountains and the smallest over the central parts of the Fennoscandian ice sheet.

3.2 Downscaling of temperature

3.2.1 Evaluation of GAMs for temperature

Equations (4) and (5) resulted in the skills presented in Table 2. For the temperature GAMs the correlations were high, explaining 96–99 % of the spatial variance of mean temperature. The MAD (RMSE) of the temperature GAMs varied between 0.6 °C (0.8 °C) in April and 1.6 °C (2.2 °C) in January. The calibration ranges of the monthly and annual mean temperatures over the inspection area (Northern Europe and Western Eurasia) are also shown in Table 2. The

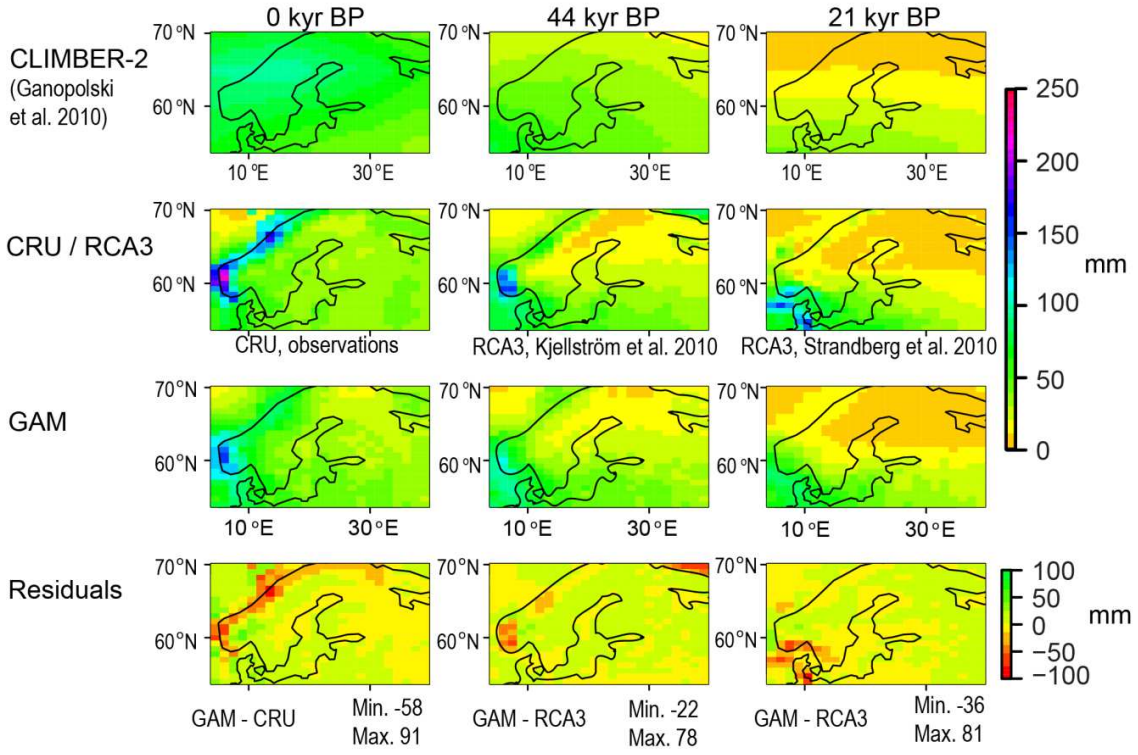

Figure 3. January mean total precipitation at 0 (left column), 44 (middle column) and 21 kyr BP (right column) as simulated by the global model CLIMBER-2 (top row), as observed during the period 1961–1990 by the CRU (first figure of the second row), as simulated by the regional model RCA3 (second and third figures of the second row), as predicted by the GAM model (third row), and as the difference between the statistical model and observation/simulation (bottom row). The unit is millimetres. The data of the first and the second rows have been bilinearly interpolated to a $1.5° \times 0.75°$ resolution.

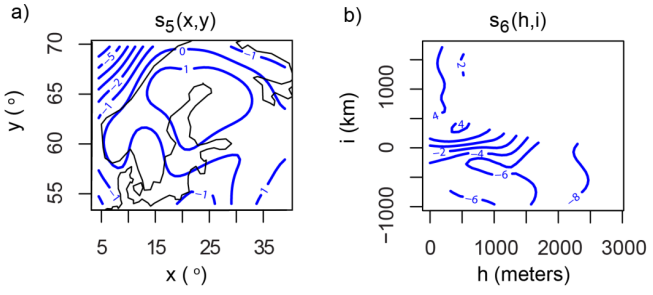

Figure 4. Components of the GAM model in Eq. (4) fitted to downscale the CLIMBER-2 July mean temperature in the recent past, MIS 3 stadial and LGM climates. The blue curves show the estimated effects of **(a)** longitude and latitude, and **(b)** elevation and distance to the nearest ice sheet (or ice cap) margin. In ice-free areas the distance was expressed by positive values and in areas covered by ice the distance was expressed by negative values.

calibration ranged from a glacial to an interglacial climate. The CLIMBER-2 temperature was used directly, i.e. linearly (see Eqs. 4, 5). With the splines for longitude and latitude some of the cold bias of CLIMBER-2 over Northern European land areas was corrected, as seen for the July temperature GAM in Fig. 4a. Over Western Eurasia, the longitude and latitude splines produced warming over Western Europe and cooling over continental Russia (not shown). In all the GAMs, the splines decreased the temperature with increasing altitude, as shown for July in Fig. 4b. In the May–September and annual GAMs (Eq. 4) the splines also decreased the temperature with increasing distance to ice-free areas, as shown for July in Fig. 4b. In the GAMs for October–April (Eq. 5) the direction of the slope also seemed to have an influence on the local temperature: the splines increased the temperature on the westerly slopes and decreased it on the easterly slopes (not shown).

Table S1.2 in the Supplement 1 shows the skills of the monthly GAMs in predicting other months' temperatures over Northern Europe. For the months May–September the correlations and errors are in the same range as for the fitting data (Table 2), suggesting that these GAMs do make reasonable predictions with other than just calibration data. For the temperature GAMs of the months October–April, the month-to-month test gave higher errors than for the calibration data (Table 2), suggesting that these GAMs are not suitable for downscaling other months than the fitted one.

Table 2. Skills and calibration ranges of the monthly mean temperature GAMs for the Northern European area, and of the annual mean temperature (∗) GAM for the Western Eurasian area. Same abbreviations as in Table 1.

Month	GCV	%ED (%)	Cor	RMSE (°C)	MAD (°C)	Min. range (°C)	Max. range (°C)
January	3.96	98	0.99	2.19	1.62	−37.3	−12.6
February	3.42	98	0.99	2.05	1.47	−36.9	−11.6
March	1.51	99	0.99	1.36	0.97	−32.1	−8.1
April	0.52	99	0.99	0.79	0.57	−23.4	−2.6
May	1.74	96	0.98	1.46	1.09	−13.9	3.8
June	1.96	96	0.98	1.53	1.15	−5.4	10.2
July	1.98	97	0.98	1.55	1.11	−0.1	13.0
August	2.10	96	0.98	1.58	1.19	−1.6	12.1
September	2.20	97	0.98	1.64	1.22	−8.5	7.7
October	1.63	99	0.99	1.42	1.04	−17.6	1.6
November	2.69	99	0.99	1.82	1.32	−26.7	−4.8
December	3.53	98	0.99	2.07	1.54	−33.9	−10.0
Annual*	2.2	98	0.99	1.8	1.3	−9.2	3.3

3.2.2 Comparison of downscaled temperatures with observations

In Fig. 5 the bilinearly interpolated CLIMBER-2 July temperatures are in the top row, the CRU-observed July temperature is in the first panel of the second row, the RCA3-simulated July temperatures are in the second and third panels of the second row, the GAM-downscaled July temperatures are in the third row, and the differences between the CRU/RCA3 data and the GAM-downscaled temperatures are in the bottom row. By comparing the second and third rows in Fig. 5, we see that the GAM reproduced well the spatial distribution of July temperature over Northern Europe, e.g. the Scandinavian mountains and the strong temperature gradient near the ice sheet margin were well brought out by the GAM. The GAM seemed to work especially well on ice-free land areas that were not in the close vicinity of ice sheets or mountains, the residuals being typically 0–2 °C. There were, however, also some differences, e.g. the downscaling of the recent past climate over the Scandinavian mountains was somewhat too cold, and the temperature gradient on the eastern margin of the ice sheet in the downscaled GAMs was less steep than in the RCA3 simulation.

Figure 6 shows pollen-based reconstructions of annual mean temperature, the bilinear interpolation of the annual mean surface temperature simulated by CLIMBER-2, and the GAM-downscaled annual mean temperature at two locations in Northern Europe during the Holocene. The pollen-based reconstructions were produced with a weighted averaging partial least squares regression and calibration technique. Bootstrapping-based root-mean-square errors of prediction for the reconstructed values are sample-specific and vary from 1.0 to 1.5 °C (Birks et al., 2010). When proxy-based reconstructions are used for comparison with model results, it must be borne in mind that the output of such reconstructions have been shown to be sensitive to the spatial extent of the calibration data sets used (Salonen et al., 2013). However, while the absolute values are highly sensitive to the climatic characteristics of the calibration data set, the shapes of the relative palaeotemperature curves seem comparatively robust, as the curve shapes mostly remain similar as the calibration data is spatially shifted (Salonen et al., 2013). Hence, it is important to note that the shapes of the reconstructed temperature curves are generally consistent with those based on GAM downscaling for the Holocene (Fig. 6). As for the absolute values, the MAD of the reconstructions and the bilinear interpolation of the CLIMBER-2 surface temperature was larger than 3.8 °C and the RMSE > 3.9 °C. The MAD of the reconstructions of the GAM output was less than 1.4 °C and the RMSE < 1.6 °C. The difference between GAM_Western Eurasia and GAM_Northern Europe is due to the difference of the driving model; that is, CCSM4 versus RCA3.

Based on Fig. 6, it seems at first that the GAM is simply a bias correction of CLIMBER simulation towards the palaeoreconstructions. The bias of the reconstructions is however unknown and thus we cannot definitely say that the GAM is an improvement over CLIMBER only because it is closer to the reconstructions (we of course hope it is, but we cannot say). Bearing in mind that the reconstructions are independent of the GAM and CLIMBER and the climate model data from CCSM4 and RCA3 are only one depiction of climate, it is nevertheless satisfying to see that the GAM is in proximity of observed climate. If we believe in the reconstructions, then we can conclude that the additional information provided to the GAM, on top of the CLIMBER simulation, is adequate and very general since the locations of the reconstructions are completely random for the GAM, and we have all reasons to further believe that the GAM would fit reconstructions in

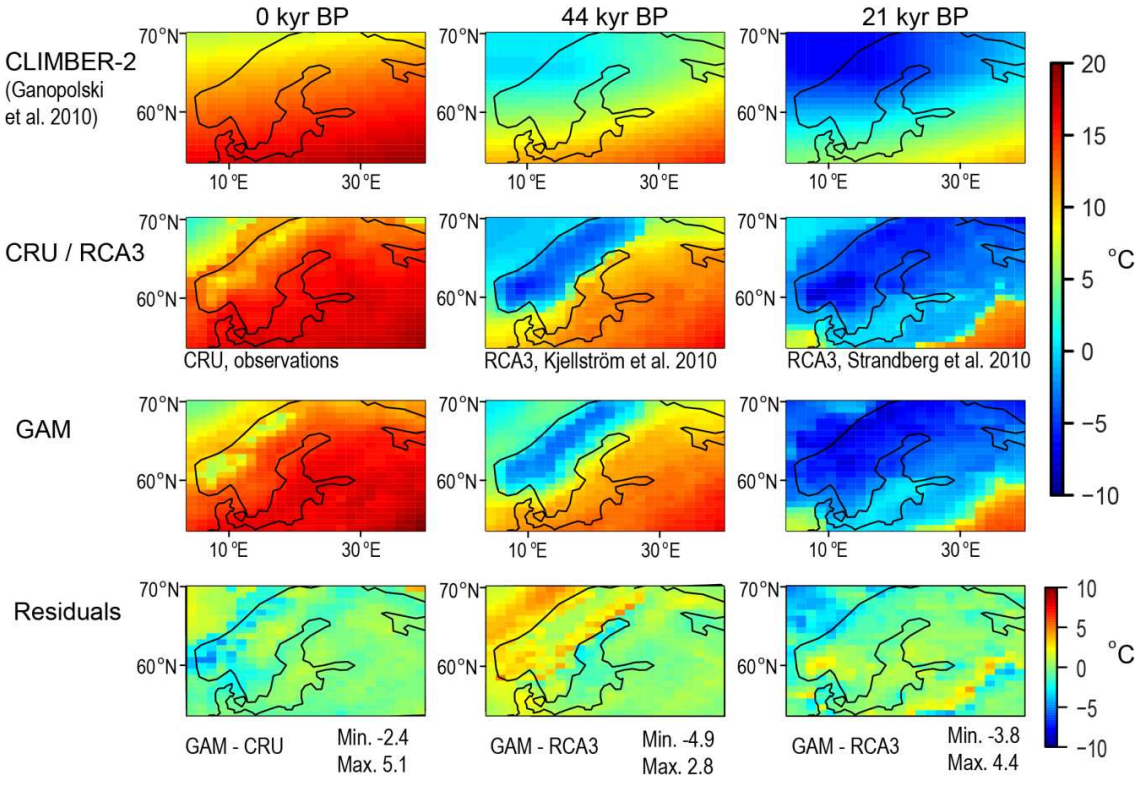

Figure 5. July mean temperature at 0 (left column), 44 (middle column) and 21 kyr BP (right column) as simulated by the global model CLIMBER-2 (top row), as observed during the period 1961–1990 by the CRU (first panel of the second row), as simulated by the regional model RCA3 (second and third panels of the second row), as predicted by the GAM model (third row), and as the difference between the statistical model and observation/simulation (bottom row). The unit is degrees Celsius. The data of the first and the second rows have been bilinearly interpolated to a $1.5° \times 0.75°$ resolution.

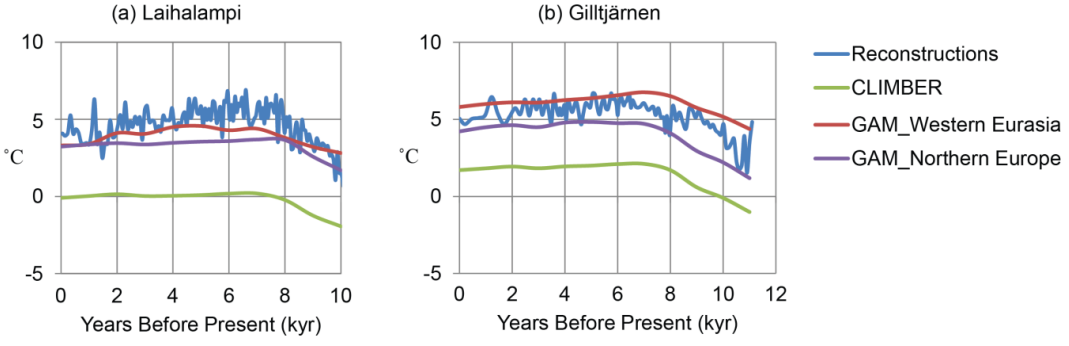

Figure 6. Comparison of simulation data and pollen-based reconstructions of annual mean temperatures from **(a)** Laihalampi in Finland and **(b)** Gilltjärnen in Sweden. The blue curves represent reconstructions by Heikkilä and Seppä (2003) for Laihalampi, and by Antonsson et al. (2006) for Gilltjärnen. The green contours are interpolations from CLIMBER-2 simulations by Ganopolski et al. (2010). The GAM_Western Eurasia (red curves) and GAM_Northern Europe (purple curves) data were downscaled by GAM models from the CLIMBER-2 data in this paper.

other locations as well. Unfortunately, there are extremely few sites with such data available.

4 Discussion and conclusions

We found that with an apt selection of predictor parameters, GAMs can be useful to statistically downscale low-resolution CLIMBER-2 EMIC simulation data ranging from a glacial to an interglacial climate. For calibrating the GAMs, we utilized

both physical predictors – such as the large-scale precipitation and temperature data of CLIMBER-2 – and geographical predictors – such as elevation, latitude, longitude, the direction of the steepest slope, the angle of the steepest slope, and the shortest distance to the ice sheet margin. The final selection of predictors was based on the statistical skill of each GAM.

The GAMs fitted for temperature (precipitation) were able to approximately reproduce the observed and modelled temperature (precipitation) fields. The mean errors in the mean temperature (total precipitation) were of an order of magnitude of 1–2 °C (10–20 mm month^{-1}) as the climate varied from glacial to interglacial over Europe. These error estimates are at the same scale as in the GAMs by Vrac et al. (2007). However, the calibration ranges in Vrac et al. (2007) covered only an interglacial climate, whereas our fitting included a glacial climate as well. Though some of the detailed spatial features were not fully captured in regions with large spatial variation of topography, the main spatial patterns were relatively well captured by the GAMs taking into account the very coarse resolution of downscaled parameters.

In Supplement 1 the monthly GAMs were used to predict other months. The skills of precipitation GAMs and temperature GAMs for months May–September in predicting other months were in the same range as for the calibration months. For the October–April temperature GAMs the errors were high, suggesting that these GAMs are not capable of downscaling outside of their calibration range. Also, Vrac et al. (2007) found that some of their GAMs calibrated over Western Europe were not valid outside their calibration domain (over North America or Northern Europe).

One critical assumption in the use of the statistical downscaling method is that the statistical relationship of the climate parameters used should remain stationary over time (Vrac et al., 2007). This assumption can be questioned with palaeoclimatic simulations spanning tens of thousands of years and reaching back to the last glacial with markedly different climatic boundary values than at present. For the validation of the statistical downscaling method, we compared our simulated Holocene temperatures with two pollen-based quantitative annual mean temperature reconstructions from Finland and Sweden. The results of these comparisons were generally congruent and thus support the validity of the statistical downscaling approach, at least in Northern Europe. Moreover, GAM-downscaled CLIMBER-2 temperature showed better agreement with the pollen-based reconstructions than the bilinearly interpolated CLIMBER-2 surface temperature.

The downscaling method enables the generation of high-resolution climate data from the low-resolution CLIMBER-2-SICOPOLIS simulations of the past. These data could be used in bioclimatic modelling and long-term climate assessments. The first comparison of the GAM-downscaled CLIMBER-2 temperature to palaeoclimatological reconstructions showed better agreement than the bilinearly interpolated CLIMBER-2 temperature. To further study the correspondence between these downscaled simulations, the next step will be to downscale the precipitation and temperature with the fitted GAMs over the whole last glacial cycle simulation, and compare the results with palaeoclimatological reconstructions. Finally, the GAMs could be further developed by adding predictors or additional simulation data such as the ensemble mean of the PMIP3 LGM simulations.

The Supplement related to this article is available online at doi:10.5194/cp-10-1489-2014-supplement.

Acknowledgements. We are grateful to Andrey Ganopolski from the Potsdam Institute for Climate Impact Research for providing the climate simulations of CLIMBER-2-SICOPOLIS, to Gustav Strandberg and Erik Kjellström from the Swedish Meteorological and Hydrological Institute for the regional climate simulations with the RCA3 model, the Climate Research Unit for the 1961–1990 high-resolution temperature and precipitation data, to NOAA/OAR/ESRL PSD, Boulder, USA, for the CMAP precipitation data, and to the contributors to the R project. The GAM fitting was done using the R software package "mgcv" downloaded from the website of the R project for statistical computing (http://www.r-project.org/). We acknowledge all those involved in producing the PMIP and CMIP multimodel ensembles and data syntheses, and the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and the climate modelling group at NCAR, Boulder, USA, for producing and making available their CCSM4 model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support, and led the development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We acknowledge Posiva PLC for funding the writing of this paper. This publication is contribution number 24 of the Nordic Centre of Excellence SVALI, Stability and Variations of Arctic Land Ice, funded by the Nordic Top-level Research Initiative (TRI). The research has been partly funded by the Academy of Finland (project no. 140771). Finally, are grateful to the anonymous referees for their good and constructive comments.

Edited by: V. Rath

References

Antonsson, K., Brooks, S. J., Seppä, H., Telford, R. J., and Birks, H. J. B.: Quantitative palaeotemperature records inferred from fossil pollen and chironomid assemblages from Lake Gilltjärnen, northern central Sweden, J. Quat. Sci., 21, 831–841, 2006.

Archer, D. and Ganopolski, A.: A movable trigger: Fossil fuel CO_2 and the onset of the next glaciation, Geochem. Geophy. Geosy., 6, Q05003, doi:10.1029/2004GC000891, 2005.

Berger, A.: Long-Term variations of Caloric Insolation Resulting from the Earth's orbital Elements, Quart. Res., 9, 139–167, 1978.

Berger, A. and Loutre, M. F.: An exceptionally long interglacial ahead?, Science, 297, 1287–1288, 2002.

Berger, A., Li, X. S., and Loutre, M. F.: Modeling northern hemisphere ice volume over the last 3 million years, Quat. Sci. Rev., 18, 1–11, 1999.

Berger, A., Loutre, M. F., and Crucifix, M.: The Earth's Climate in the Next hundred thousand years (100 kyr), Surv. of Geophys., 24, 117–138, 2003.

Birks, H. J. B., Heiri, O., Seppä, H., and Bjune, A. E.: Strengths and weaknesses of quantitative climate reconstructions based on Late-Quat. biological proxies, Open Ecol. J., 3, 68–110, 2010.

Braconnot, P., Harrison, S., Otto-Bliesner, O., Abe-Ouchi, A., Jungclaus, J., and Peterschmitt, J.-Y.: The Paleoclimate Modeling Intercomparison Project contribution to CMIP5, CLIVAR Exchanges, 56, 15–19, 2011.

Braconnot, P., Harrison, S., Kageyama, M., Bartlein, P., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B., and Zhao, Y.: Evaluation of climate models using palaeoclimatic data, Nat. Clim. Change, 2, 417–424, 2012.

Brady, E., Otto-Bliesner, B., Kay, J., and Rosenbloom, N.: Sensitivity to Glacial Forcing in the CCSM4, J. Climate, 36, 1901–1925, 2013.

Brandefelt, J. and Otto-Bliesner, B.: Equilibration and variability in a Last Glacial Maximum Climate simulation with CCSM3, Geophys. Res. Lett., 36, L19712, doi:10.1029/2009GL040364, 2009.

Calov, R., Ganopolski, A., Claussen, M., Petoukhov, V., and Greve, R.: Transient simulation of the last glacial inception. Part I: glacial inception as a bifurcation in the climate system, Clim. Dynam., 24, 545–561, 2005a.

Calov, R., Ganopolski, A., Claussen, M., Petoukhov, V., Claussen, M., Brovkin, V., and Greve, R.: Transient simulation of the last glacial inception, Part II: sensitivity and feedback analysis, Clim. Dynam., 24, 563–576, 2005b.

Cheng, M. and Qi, Y.: Frontal rainfall rate distribution and some conclusions on the threshold method, J. App. Meteorol., 41, 1128–1139, 2002.

Claussen, M., Mysak, L., Weaver, A., Crucifix, M., Fichefet, T., Loutre, M.F., Weber, S., Alcamo, J., Alexeev, V., Berger, A., Calov, R., Ganopolski, A., Goosse, H., Lohmann, G., Lunkeit, F., Mokhov, I., Petoukhov, V., Stone, P., and Wang, Z.: Earth system Models of Intermediate Complexity: Closing the gap in the spectrum of climate system models, Clim. Dynam., 18, 579–586, 2002.

Cochelin, A.-S. B., Mysak, L. A., and Wang, Z.: Simulation of long-term future climate changes with the green McGill paleoclimate model: The next glacial inception, Clim. Change, 79, 381–401, 2006.

Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., Chang, P., Doney, S. C., Hack, J. J., Henderson, T. B., Kiehl, J. T., Large, W. G., McKenna, D. S., Santer, B. D., and Smith, R. D.: The Community Climate System Model Version 3CCSM3), J. Climate, 19, 2122–2143, 2006.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 746–747, 2013.

Ganopolski, A., Petoukhov, V., Rahmstorf, S., Brovkin, V., Claussen, M., Eliseev, A., and Kubatzki, C.: CLIMBER-2: a climate system model of intermediate complexity, Part II: model sensitivity, Clim. Dynam., 17, 735–751, 2001.

Ganopolski, A., Calov, R., and Claussen, M.: Simulation of the last glacial cycle with a coupled climate ice-sheet model of intermediate complexity, Clim. Past, 6, 229–244, doi:10.5194/cp-6-229-2010, 2010.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne S. R., Lawrence, D. M., Neale, R. B., Rasch P. J., Vertenstein, M., Worley, P. H., Yang, Z.-L., and Zhang, M.: The Community Climate System Model Version 4, J. Climate, 24, 4973–4991, 2011.

Greve, R.: Application of a polythermal three-dimensional ice sheet model to the Greenland ice sheet: response to steady-state and transient climate scenarios, J. Climate, 10, 901–918, 1997.

Hastie, T. and Tibshirani, R.: Generalized Additive Models , monographs on statistics and applied probability, vol. 43, Chapman & Hall, New York, 1990.

Heikkilä, M. and Seppä, H.: A 11 000 yr palaeotemperature reconstruction from the southern boreal zone in Finland, Quat. Sci. Rev., 22, 541–554, 2003.

Jones, P. D., New, M., Parker, D. E., Martin, S., and Rigor, I. G.: Surface air temperature and its changes over the past 150 years, Rev. Geophys., 37, 173–199, 1999.

Kageyama, M., Laîné, A., Abe-Ouchi, A., Braconnot, P., Cortijo, E., Crucifix, M., de Vernal, A., Guiot, J., Hewitt, C. D., Kitoh, A., Kucera, M., Marti, O., Ohgaito, R., Otto-Bliesner, B., Peltier, W. R., Rosell-Melé, A., Vettoretti, G., Weber, S. L., Yu, Y., and MARGO Project members: Last Glacial Maximum temperatures over the North Atlantic, Europe and western Siberia: a comparison between PMIP models, MARGO sea-surface temperatures and pollen-based reconstructions, Quat. Sci. Rev., 25, 2082–2102, 2006.

Kiehl, J. T., Shields, C. A., Hack, J. J., and Collins, W. D.: The Climate Sensitivity of the Community Climate System Model Version 3 (CCSM3), J. Climate, 19, 2584–2596, 2006.

Kim, Y. J. and Gu, C.: Smoothing spline Gaussian regression: more scalable computation via efficient approximation. Journal of the Royal Statistical Society, Series B, 66, 337–356, 2004.

Kistler, R., Kalnay, E., Collins,W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki,W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, Bull. Amer. Meteorol. Soc., 82, 247–26. 2001.

Kjellström, E., Bärring, L., Gollvik, S., Hansson, U., Jones, C., Samuelsson, P., Rummukainen, M., Ullerstig, A., Willén, U., and Wyser, K.: A 140-year simulation of European climate with the new version of the Rossby Centre regional atmospheric climate model (RCA3), Reports Meteorology and Climatology, 108, SMHI, SE-60176 Norrköping, Sweden, 54 pp., 2005.

Kjellström, E., Brandefelt, J., Näslund, J. O., Smith, B., Strandberg, G., Voelker, A. H. L., and Wohlfarth, B.: Simulated climate con-

ditions in Fennoscandia during a MIS 3 stadial, Boreas, 39, 436–456, 2010.

Liu, Z., Otto-Bliesner, B. L., He, F., Brady, E. C., Tomas, R., Clark, P. U., Carlson, A. E., Lynch-Stieglitz, J., Curry, W., Brook, E., Erickson, D., Jacob, R., Kutzbach, J., and Cheng, J.: Transient Simulation of Last Deglaciation with a New Mechanism for Bølling-Allerød Warming, Science, 325, 310–314, 2009.

Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol., 25, 693–712, 2005.

Otto-Bliesner, B.L., Brady, E.C., Clauzet, G., Tomas, R., Levis, S., and Kothavala, Z.: Last glacial maximum and Holocene climate in CCSM3, J. Climate, 19, 2526–2544, 2006.

Peltier, W. R.: Global glacial isostacy and the surface of the ice-age circulation, Earth: The ICE-5G (VM2) model and GRACE, Annu. Rev. Earth Pl. Sc., 32, 111–149, 2004.

Petoukhov, V., Ganopolski A., Brovkin V., Claussen M., Eliseev A., Kubatzki, and Rahmstorf, S.: CLIMBER-2: A climate system model of intermediate complexity. Part I: Model description and performance for present climate, Clim. Dynam., 16, 1–17, 2000.

Petoukhov, V., Claussen, M., Berger, A., Crucifix, M., Eby, M., Eliseev, A. V., Fichefet, T., Ganopolski, A., Goosse, H., Kamonkovich, I., Mokhov, I. I., Montoya, M., Mysak, L. A., Sokolov, A., Stone, P., Wang, Z., and Weaver, A. J.: EMIC Intercomparison Project (EMIP-CO_2): comparative analysis of EMIC simulations of climate, and of equilibrium and transient responses to atmospheric CO_2 doubling, Clim. Dynam., 25, 363–385, 2005.

Racherla, P. N., Shindell, D. T., and Faluvegi, G.: The added value to global model projections of climate change by dynamical downscaling: A case study over the continental US using the GISS-ModelE2 and WRF models, J. Geophys. Res., 117, D20118, doi:10.1029/2012JD018091, 2012:

Renssen, H., Goosse, H., Fichefet, T., Brovkin, V., Driesschaert, E., and Wolk, F.: Simulating the Holocene climate evolution at northern high latitudes using a coupled atmosphere-sea ice-ocean-vegetation model, Clim. Dynam., 24, 23–43, 2005.

Salonen, J. S., Seppä, H., and Birks, H. J. B.: The effect of calibration data-set selection on quantitative palaeoclimate reconstructions, Holocene, 23, 1650–1654, 2013.

Samuelsson, P., Jones, C., Willén, U., Ullerstig, A., Gollvik, S., Hansson, U., Kjellström, E., Nikulin, G., and Wyser, K.: The Rossby Centre Regional Climate Model RCA3: Model description and performance, Tellus A, 63, 4–23, 2011.

Singarayer, J. S. and Valdes, P. J.: High-latitude climate sensitivity to ice-sheet forcing over the last 120 kyr, Quat. Sci. Rev., 29, 43–55, 2010.

Schneider von Deimling, T., Held, H., Ganopolski, A., and Rahmstorf, S.: Climate sensitivity estimated from ensemble simulations of glacial climate, Clim. Dynam., 27, 149–163, 2006.

Smith, R. S. and Gregory, J.: The last glacial cycle: transient simulations with an AOGCM. Clim. Dynam., 38, 1545–1559, 2012.

Strandberg, G., Brandefelt, J., Kjellström, E., and Smith, B.: High resolution regional simulation of Last Glacial Maximum climate in Europe, Tellus A, 63, 107–125, 2011.

Uppala, S., Kallberg, P., Simmons, A., Andrae, U., Bechtold, V., Fiorino, M., Gibson, J., Haseler, J., Hernandez, A., Kelly, G., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R., Andersson, E., Arpe, K., Balmaseda, M., Beljaars, A., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, E., Hoskins, B., Isaksen, L., Janssen, P., Jenne, R., McNally, A., Mahfouf, J., Morcrette, J., Rayner, N., Saunders, R., Simon, P., Sterl, A., Trenberth, K., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Q. J. R. Meteorol. Soc., 131, 2961–3012, 2005.

Vrac, M., Marbaix, P., Paillard, D., and Naveau, P.: Non-linear statistical downscaling of present and LGM precipitation and temperatures over Europe, Clim. Past, 3, 669–682, doi:10.5194/cp-3-669-2007, 2007.

Wang, Z. and Mysak, L.: Simulation of the last glacial inception and rapid ice sheet growth in the McGill Paleoclimate Model, Geophys. Res. Lett., 29, 2102, doi:10.1029/2002GL015120, 2002.

Wood, S. N.: Generalized additive Models: An Introduction with R, CRC/Chapman and Hall, Boca Raton, Florida, 2006.

Xie, P. and Arkin, P. A.: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, B. Am. Meteorol. Soc., 78, 2539–2558, 1997.

Supplement of Clim. Past, 10, 1489–1500, 2014
http://www.clim-past.net/10/1489/2014/
doi:10.5194/cp-10-1489-2014-supplement
© Author(s) 2014. CC Attribution 3.0 License.

Climate of the Past
Open Access

Supplement of

Statistical downscaling of a climate simulation of the last glacial cycle: temperature and precipitation over Northern Europe

N. Korhonen et al.

Correspondence to: N. Korhonen (natalia.korhonen@fmi.fi)

SUPPLEMENT to manuscript
"Statistical downscaling of a climate simulation of the last glacial cycle: temperature and precipitation over Northern Europe" by N. Korhonen, A. Venäläinen, H. Seppä, H. Järvinen

To widen the validation of the statistical downscaling we tested the monthly GAMs in predicting other months' temperature or precipitation over Northern Europe. This was done by predicting with the GAM calibrated by the previous month's precipitation the precipitation by CLIMBER (P_{CLI}). The prediction, e.g. in the case of the February precipitation GAM run by January CLIMBER precipitation data was compared to the CRU, simulated 44 kyr BP (Kjellström et al. 2010) and simulated LGM (Strandberg et al., 2011) January precipitation data (reference data). We calculated the Pearson correlation, RMSE and MAD between the predicted and reference data. The results are in Tables S1.1 and S1.2 below.

For the precipitation GAMs and the temperature GAMs for the months May-September the correlations and errors are in the same range as for the fitting data (compare to Table 1 and 2 in the paper), suggesting that these GAMs do predict reasonable also for other than just calibration data.
For the temperature GAMs of the months October-April, the month to month test gave higher errors than for the calibration data (Table 2 in the paper) suggesting that these GAMs have problems working outside their calibration domain.

Table S1.1. Validating monthly precipitation GAMs on other months.

Calibration month	Validation month	Cor	RMSE (mm month^{-1})	MAD (mm month^{-1})
January	February	0.84	13	8
February	January	0.88	16	9
March	April	0.61	18	13
April	March	0.65	24	17
May	June	0.61	18	14
June	May	0.59	17	12
July	August	0.68	19	13
August	July	0.75	16	12
September	October	0.73	25	18
October	September	0.69	26	17
November	December	0.82	20	14
December	November	0.79	23	14

Table S1.2. Validating monthly temperature GAMs on other months.

Calibration month	Validation month	Cor	RMSE (˚C)	MAD (˚C)
January	February	0.90	7.98	5.5
February	January	0.97	4.02	3.34
March	April	0.96	8.27	6.93
April	March	0.92	10.33	8.41
May	June	0.97	1.96	1.6
June	May	0.97	1.92	1.56
July	August	0.98	1.99	1.63
August	July	0.98	1.95	1.59
September	August	0.96	2.73	2.33
October	September	0.97	4.42	3.2
November	December	0.93	7.06	5.45

December	November	0.97	5.13	3.98

SUPPLEMENTARY MATERIAL 2 to manuscript
"Statistical downscaling of a climate simulation of the last glacial cycle: temperature and precipitation over Northern Europe" by N. Korhonen, A. Venäläinen, H. Seppä, H. Järvinen

The annual GAMs for Western Eurasia (calibrated only by recent past and LGM data) were used for predicting annual mean temperature and precipitation for 44 kyr BP by CLIMBER-2-SICOPOLIS 44 kyr BP data from the last glacial cycle simulation by Ganopolski et al. (2010). The predictions are plotted below, in Figures S2.1 and S2.2, together with output of CLIMBER-2 and RCA3 regional climate model simulation output by Kjellström et al. (2010) representing 44 kyr BP annual mean temperature and precipitation.

Figures S2.1 b) and c) show that over Eastern Europe the GAM predicts similar temperature spatial patterns as the RCA3, whereas over Western Europe, there are more differences in details. The differences between RCA3 and GAM temperature over Finland mainly result from the difference of ice sheet extent in SICOPOLIS and RCA3: the Fennoscandian ice sheet in SICOPOLIS reaches Central Finland during 44 kyr BP, whereas in the RCA3 simulation the ice sheet reaches only some parts of Northern Finland (Figure 1 in paper).

The GAM predictions of precipitation in Fig. S2.2 c) are similar to the RCA3 simulation output (Fig. S2.2 b) in many parts of Eastern and Central Europe, however, not predicting as high precipitation amounts in mountain regions.

a) CLIMBER-2
Ganopolski et al. 2010

b) RCA3
Kjellstöm et al. 2010

c) GAM Western Eurasia

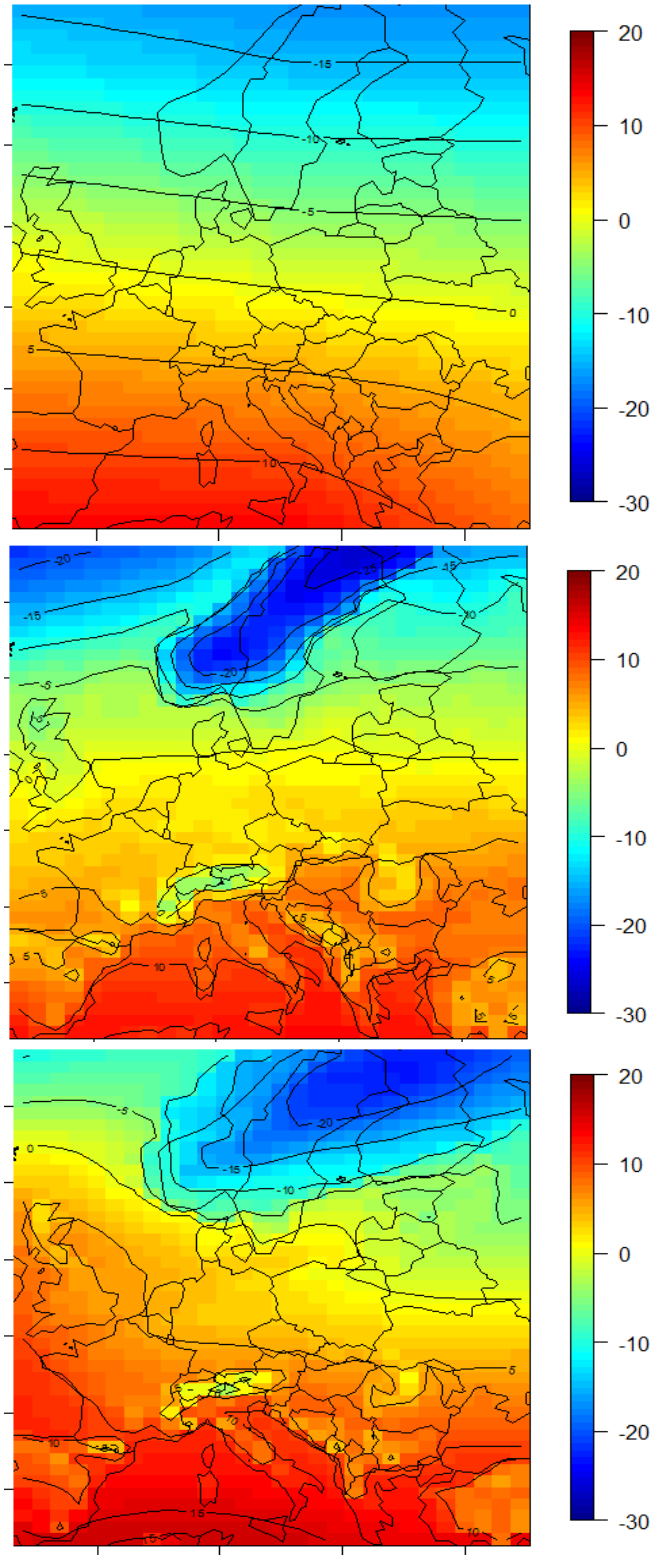

Fig. S2.1 Annual mean temperature at 44 kyr BP, a) as simulated by the global model CLIMBER-2 (Ganopolski et al. 2010), b) as simulated by the regional model RCA3 (Kjellström et al. 2010), and c) as predicted by the GAM model "Western Eurasia". Unit: Celsius. The data of CLIMBER-2 and RCA3 have been bi-linearly interpolated on to a 1.5º × 0.75º resolution.

a) CLIMBER-2
Ganopolski et al. 2010

b) RCA3
Kjellstöm et al. 2010

c) GAM Western Eurasia

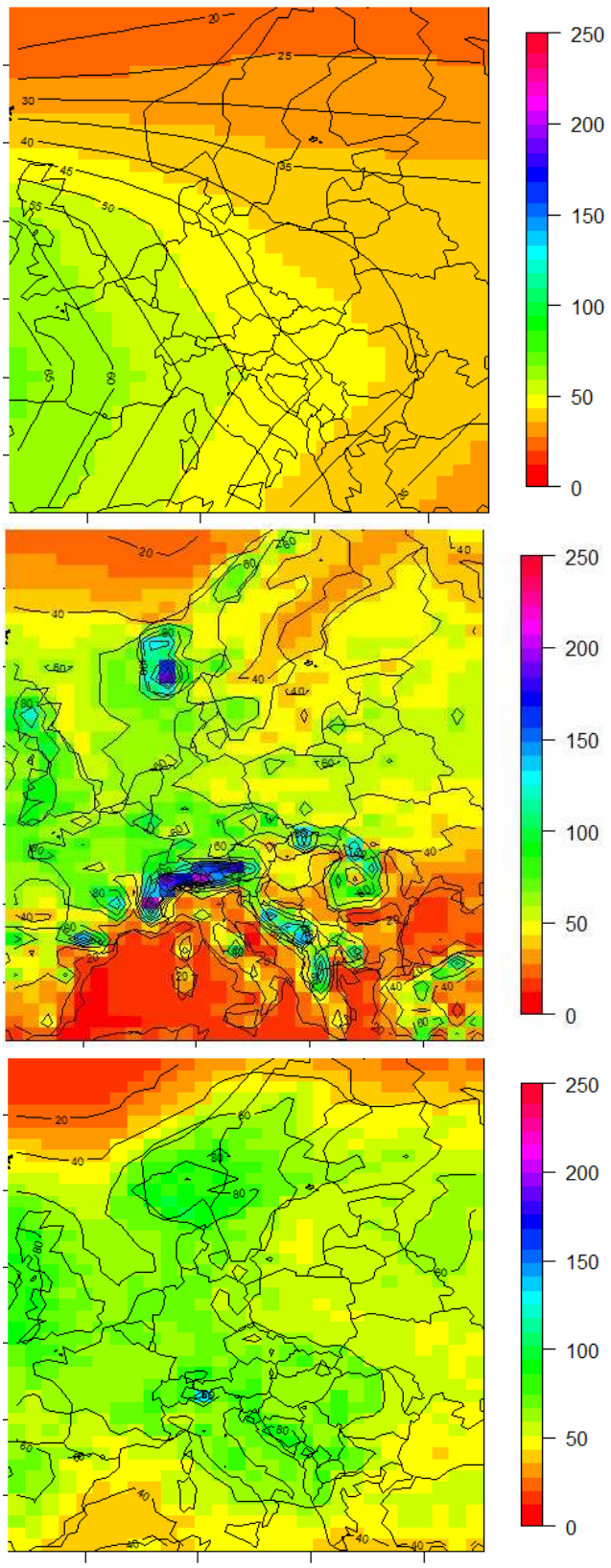

Fig. S2.2 a) Annual mean total precipitation at 44 kyr BP, a) as simulated by the global model CLIMBER-2 (Ganopolski et al. 2010), b) as simulated by the regional model RCA3 (Kjellström et al. 2010), and b) as predicted by the GAM model "Western Eurasia". Unit: mm month^{-1}. The data of CLIMBER-2 and RCA3 have been bi-linearly interpolated on to a 1.5º × 0.75º resolution.

Paper IV

Reprinted, with permission from
Proceedings of the National Academy of Sciences of the United States of America,
112: 8232–8237, 2015.

Human population dynamics in Europe over the Last Glacial Maximum

Miikka Tallavaara[a,1], Miska Luoto[b], Natalia Korhonen[c,d], Heikki Järvinen[d], and Heikki Seppä[b]

[a]Department of Philosophy, History, Culture and Art Studies, University of Helsinki, FI-00014 Helsinki, Finland; [b]Department of Geosciences and Geography, University of Helsinki, FI-00014 Helsinki, Finland; [c]Climate Change Research Unit, Finnish Meteorological Institute, FI-00101 Helsinki, Finland; and [d]Department of Physics, University of Helsinki, FI-00014 Helsinki, Finland

Edited by Jean-Pierre Bocquet-Appel, Ecole Pratique des Hautes Etudes, Paris, France, and accepted by the Editorial Board May 21, 2015 (received for review February 25, 2015)

The severe cooling and the expansion of the ice sheets during the Last Glacial Maximum (LGM), 27,000–19,000 y ago (27–19 ky ago) had a major impact on plant and animal populations, including humans. Changes in human population size and range have affected our genetic evolution, and recent modeling efforts have reaffirmed the importance of population dynamics in cultural and linguistic evolution, as well. However, in the absence of historical records, estimating past population levels has remained difficult. Here we show that it is possible to model spatially explicit human population dynamics from the pre-LGM at 30 ky ago through the LGM to the Late Glacial in Europe by using climate envelope modeling tools and modern ethnographic datasets to construct a population calibration model. The simulated range and size of the human population correspond significantly with spatiotemporal patterns in the archaeological data, suggesting that climate was a major driver of population dynamics 30–13 ky ago. The simulated population size declined from about 330,000 people at 30 ky ago to a minimum of 130,000 people at 23 ky ago. The Late Glacial population growth was fastest during Greenland interstadial 1, and by 13 ky ago, there were almost 410,000 people in Europe. Even during the coldest part of the LGM, the climatically suitable area for human habitation remained unfragmented and covered 36% of Europe.

hunter-gatherers | demography | niche modeling | climate change | Paleolithic

Growing populations of anatomically and behaviorally modern humans have been partly responsible for past ecosystem changes such as the extinctions of Pleistocene megafauna and Neanderthal humans (1, 2). In addition to the destiny of other species, human population size also influences our own cultural and genetic evolution. Large pools of interacting individuals can create and maintain adaptive skills, as well as phonological variation, more effectively than small populations, and they are also capable of faster cumulative cultural evolution (3–5). A decrease in population size may even result in a loss of complex cultural traits (6). The effects of population size on cultural variation are thus roughly similar to the effects of population size on genetic variation (7).

The study of the role of human population size in cultural and genetic evolution and past ecosystem changes necessitates estimates of population dynamics extending far beyond historical times. The archaeological record illustrates patterns of human population range and size dynamics (8–10), but it does not offer quantitative population size data. Archaeological reconstructions of population dynamics are also bound to the regions and time periods that offer a sufficiently rich archaeological record. In addition to archaeological data, information on past population patterns can be inferred from genetic data using skyline-plot methods (11) and pairwise or multiple sequentially Markovian coalescent analyzes (12, 13). However, these methods depend on estimates of DNA mutation rate and molecular clock calibrations, which are still debated (14, 15) and imprecise, leading to poor temporal resolution. Furthermore, these methods track changes in effective population size that does not have a straightforward relationship with the actual census population size (16). Together with poor resolution, this makes it extremely difficult to meaningfully compare DNA-based population reconstructions with the records of cultural or environmental changes.

Here, we take a different approach and model human population size and range dynamics in the last glacial Europe independently of archaeological and genetic data. Ethnographic studies have found a link between climate and the diet, mobility, and territory size of hunter-gatherers (17–20). We hypothesize that correlation exists also between climate and hunter-gatherer population density. We take advantage of this potential climate connection and use an approach made possible by recent developments in climate envelope modeling.

Climate envelope or niche models use associations between aspects of climate and the occurrences of species to estimate the conditions that are suitable for maintaining viable populations (21–23). By using information on how the climate affects modern hunter-gatherer population densities, this framework allows us to evaluate climatic suitability for humans and simulate their potential distribution and abundance under the changing climatic conditions of the last glacial, thus overcoming the above-mentioned limitations of approaches using only archaeological or genetic data. We use ethnographic data on terrestrially adapted mobile hunter-gatherers and their climatic space (24) (Dataset S1) to construct a calibration model that predicts hunter-gatherer presence and population density by three climatic predictors: potential evapotranspiration and water balance, both of which exert strong

Significance

Despite its importance for understanding genetic, cultural, and linguistic evolution, prehistoric human population history has remained difficult to reconstruct. We show that the dynamics of the human population in Europe from 30,000 to 13,000 y ago can be simulated using ethnographic and paleoclimate data within the climate envelope modeling approach. Correspondence between the population simulation and archaeological data suggests that population dynamics were indeed driven by major climate fluctuations, with population size varying between 130,000 and 410,000 people. Although climate has been an important determinant of human population dynamics, the climatic conditions during the last glacial were not as harsh as is often presented, because even during the coldest phases, the climatically suitable area for humans covered 36% of Europe.

Author contributions: M.T., M.L., and H.S. designed research; M.T., M.L., N.K., H.J., and H.S. performed research; M.T. and M.L. contributed new reagents/analytic tools; M.T., M.L., N.K., and H.J. analyzed data; and M.T., M.L., N.K., H.J., and H.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.-P.B.-A. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: miikka.tallavaara@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1503784112/-/DCSupplemental.

influence on ecosystem productivity and species richness, and the mean temperature of the coldest month that affects wintering conditions, such as winter mortality (25, 26). This model is extrapolated over the European landscape for 30–13 ky ago using climate predictor values obtained by statistical downscaling of the CLIMBER-2 climate model simulation data (27, 28) (Dataset S2). The period in question was chosen because it extends from the end of the Marine Isotope Stage 3 (MIS-3) to the Last Glacial Termination and includes the coldest phase and the largest ice sheet extent of the last glaciation (29).

In practice, estimates of absolute prehistoric population size or density cannot be truly verified with any existing data. Because our model is not archaeologically informed, it is, however, possible to use the dataset of archaeological radiocarbon dates (30) (Dataset S3) to evaluate the simulated spatial and temporal patterns and, in that sense, the realism of our simulation. Such archaeological data are increasingly used as a proxy in studies of prehistoric human population dynamics (8–10, 31, 32).

Results

Fig. 1 shows that the temporal patterns in the simulated population size and archaeological population proxy are remarkably consistent (r_P = 0.84, P < 0.00002). Both show relatively high late-MIS-3 population size levels, a decline toward the Last Glacial Maximum (LGM) minimum, and a rapid growth during the Late Glacial. The simulation suggests that the human population size in Europe was about 330,000 at 30 ky ago, 130,000 during its minimum at 23 ky ago, and almost 410,000 at 13 ky ago, during the Greenland interstadial 1. The mean population density in the inhabited area varied between 2.8 and 5.1 persons per 100 km^2.

The simulated spatial pattern of human population (Figs. 2 and 3) indicates a population contraction starting in line with the ice sheet expansion at 27 ky ago. During the peak LGM, the northern limit of contiguous population in Europe extended from central France to lowlands in southern Germany and to the southern parts of modern Ukraine and European Russia (Fig. 2). Thus, there was an uninhabited zone about 500 km wide between the ice sheet and the northern limit of the human population. However, our simulation suggests that the continuously suitable and inhabited area between 30 and 13 ky ago covered 36% of the European land area even during the coldest LGM, stretching to the north of the Alps (Fig. 3), a result supported also by an emerging archaeological picture (33). In addition, the simulation shows a persistent southwest-northeast gradient of decreasing population densities, with the densest populations throughout the LGM in the Iberian Peninsula and the Mediterranean region (Figs. 2 and 3). The post-LGM recolonization of the continent began at 19 ky ago.

These spatial dynamics concur with the archaeological data, although the latter show a spatially and temporally more sporadic pattern. Such a sparse pattern is most probably a result of a Wallacean shortfall-like effect of incomplete information on species distribution. Although Wallacean shortfall is true for current plant and animal species, paleontological and archaeological records provide obviously even more incomplete and coarse reflection of true ranges (23).

There are, nevertheless, two instances where the simulated range and density of the human population deviate from the distribution of archaeological data. First, in northern Russia, the archaeological record indicates occasional presence of humans much farther north than our simulation suggests. These anomalies may represent human populations whose climatic tolerance differed from that of the modern hunter-gatherer populations used in the calibration model, because it has been suggested that the archaeological lithic assemblage of Byzovaya site in the Polar Urals was produced by Neanderthals (34). The presence of Neanderthals is controversial (35), however, and the anomalies continue sporadically throughout the LGM, when Neanderthals

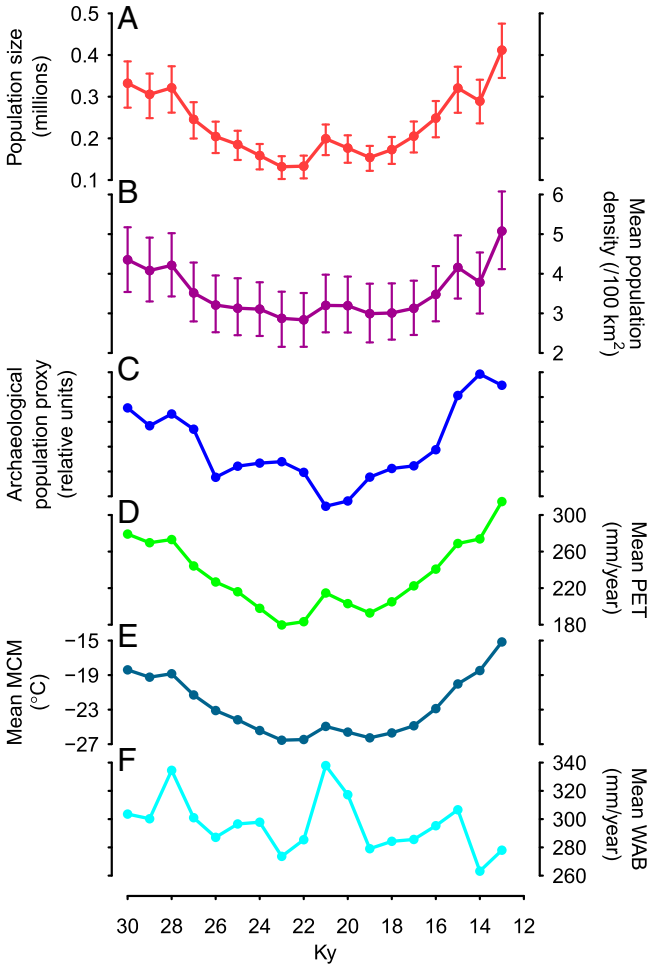

Fig. 1. Comparisons between simulated hunter-gatherer population size and density, the archaeological population proxy, and paleoclimatic simulations between 30 and 13 ky ago in Europe. (*A*) Simulated human population size in Europe. Error bars show the resampling-based confidence limits (95%). (*B*) Simulated mean density in the inhabited area of Europe. Error bars show the resampling-based confidence limits (95%). (*C*) Archaeological population size proxy based on the taphonomically corrected number of dates. (*D*) European mean of simulated potential evapotranspiration. (*E*) European mean of simulated mean temperature of the coldest month. (*F*) European mean of simulated water balance. *D*–*F* are based on the downscaling from the CLIMBER-2 climate model.

are assumed to have already been extinct. Nonetheless, our results allow for the possibility that the late MIS-3 populations in northern Russia were biologically or behaviorally different from later humans.

Second, whereas our model simulates high population densities in the Mediterranean region, the density of archaeological data in the region is relatively low throughout the study period. This difference does not relate to the properties of the climate data used in the simulation, because the LGM snapshot population simulations based on state-of-the-art general circulation model data (36–38) show the same pattern (Fig. S1). This similarity of the patterns strongly suggests that the Iberian Peninsula and the Mediterranean region have indeed been climatically the most suitable areas for hunter-gatherers throughout the LGM.

It is possible that some nonclimatic factors made the region less suitable for humans, which would explain the difference between the simulation based on the hunter-gatherer climatic envelope and the archaeological data. For example, the climatically highly

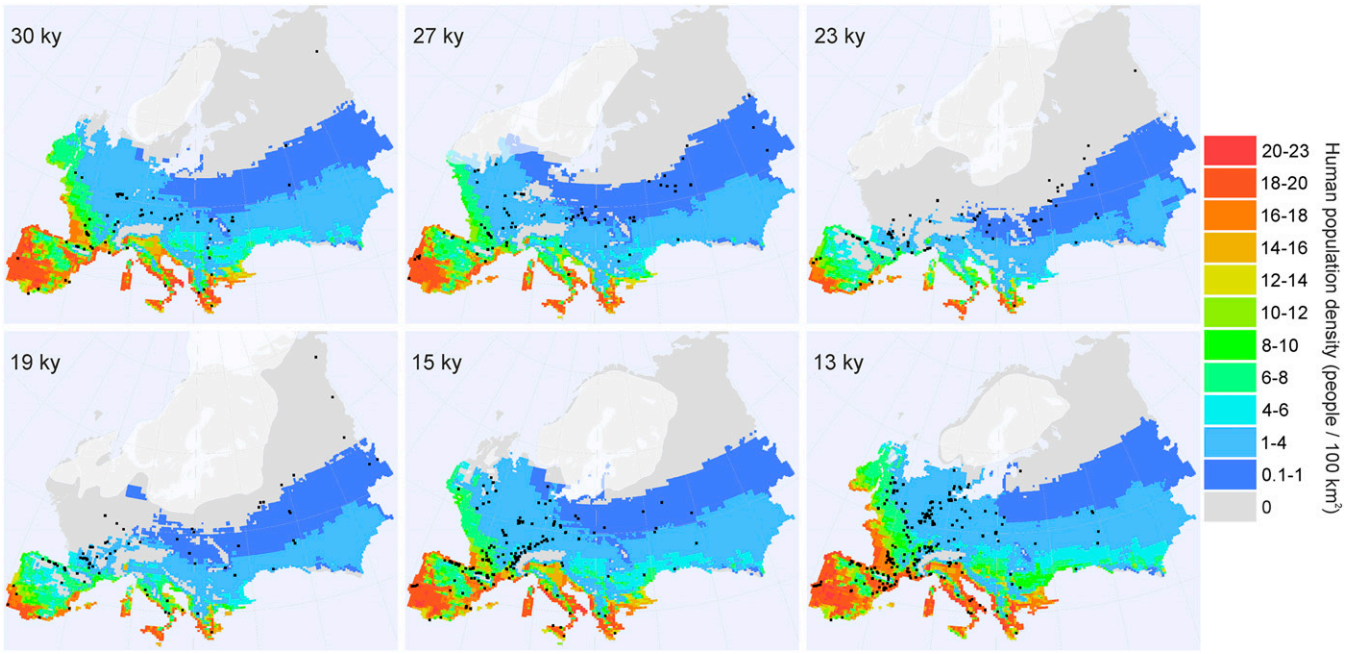

Fig. 2. Simulated human population range and density compared with the spatial distribution of archaeological sites during six time intervals from 30 to 13 ky ago. Archaeological sites are indicated with black dots and in each time slice they represent sites dated within 1,000-y bins.

suitable, but relatively small, island of Sardinia may not have been attractive for terrestrially adapted hunter-gatherer groups of the LGM Europe. On a larger scale, however, the spatial distribution of archaeological data may not adequately reflect the distribution of human population, because of the systematic differences in taphonomic processes between different parts of Europe. Due to a combination of climatic and topographic features, erosion rates are higher in the Mediterranean region than in other parts of Europe (39, 40). The high erosion leads to loss or disturbance of the sediment layers containing archaeological material, which may explain the relatively low density of radiocarbon dated sites in the Mediterranean region. For example, in a sample of 164 Middle Paleolithic sites in southern Iberia, almost 80% of the sites were found to be clearly in a secondary context (41).

Research history may also play role in the spatial variability of archaeological data. In Portugal, for instance, there were only four Upper Paleolithic sites known in the early 1960s, and the region was considered largely uninhabited (42). Sensitivity of archaeological distributions to changes in research interests is reflected by the fact that in 50 y the number of sites has multiplied manifold with such discoveries as the Côa Valley dwelling and rock art sites (43, 44). However, relatively few of these new sites have been radiocarbon dated (44) and would not show up in our archaeological proxy. It is thus likely that the discrepancy between the simulated population densities and the spatial distribution of archaeological data in the Mediterranean region is a result of combined effect of research history and erosion-induced taphonomic loss and disturbance of archaeological material. In general, the archaeological data, nevertheless, fall within the simulated range area and the northern limits of the simulation and the archaeological data correspond to each other relatively well.

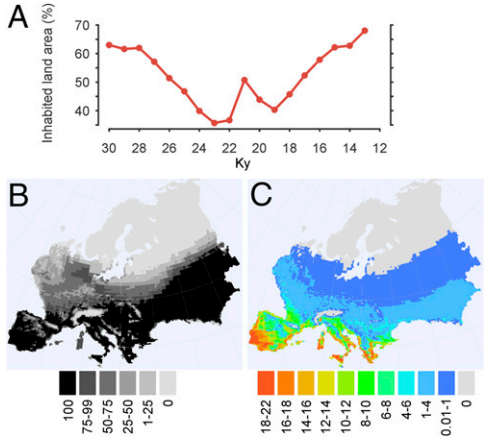

Fig. 3. Climatic suitability of Europe for human population over the LGM according to the simulation. (*A*) Changes in the percentage of potentially inhabited land area in Europe. (*B*) Percentage of time the area has potentially been inhabited between 30 and 13 ky ago. (*C*) Mean population density (people/100 km^2) between 30 and 13 ky ago.

Discussion

The overall similarity of the simulated and archaeological population patterns supports our results about the European human population changes between 30 and 13 ky ago. However, the simulated population size in LGM Europe appears extremely high compared with the results of Bocquet-Appel et al. (45), who estimated the population size to be less than 6,000 persons. There are two main reasons that lead to these considerably smaller population size estimates. First, Bocquet-Appel et al. (45) estimate the human population range from the spatial distribution of archaeological data while assuming that it adequately reflects the true range of the human population. As discussed above, this assumption is probably not valid, because archaeological remains provide an incomplete and coarse reflection of past geographical distributions of human activity. Second, compared with ethnographically known hunter-gatherer populations (17, 24), Bocquet-Appel et al. (45) use extremely low population density estimates and, even more importantly, only single estimates for each period in question, which does not take into account geographical variability in climate

and environment. As our simulation shows, this variability has a strong impact on human population densities and this result appears to be robust with respect to the choice of climate model simulation data (Fig. S1).

In addition, the simulated temporal pattern in population size and density differs from many previous reconstructions of Pleistocene dynamics of European populations, especially those based on DNA data (11–13), which show monotonic growth for the last 50 ky. However, the monotonic growth does not match the archaeological record, which shows substantial variations in human population size at high and low frequencies. These variations seem often to follow climatic variability and, as shown here, they are also simulated with the climate envelope modeling.

Our results have three important implications. First, we show that the range and size of prehistoric hunter-gatherer populations can be realistically modeled using information on modern hunter-gatherers and paleoclimatic simulations. This climate envelope modeling approach provides valuable insights into the patterns and causes of long-term human population dynamics and a necessary complement to archaeological and DNA-based methods when studying prehistoric human demography. Second, the consistency between the simulated patterns and the archaeological data are remarkable because it suggests that climatic conditions were crucial drivers of last glacial human population dynamics. This consistency also indicates that the climate envelope of the hunter-gatherers has remained relatively constant from the last glacial to present. Millennia of cultural evolution have not fundamentally changed constraints on terrestrially adapted hunter-gatherer populations posed by the climate. Third, even the harsh conditions of the LGM sustained a substantial human population in Europe, which was not fragmented to totally isolated refugia. The continuous range would have facilitated a flow of genes and cultural information between the western and eastern parts of the continent, which, in turn, has implications for understanding genetic diversity and cultural evolution in Europe.

Materials and Methods

Constructing the Calibration Model. We used an ethnographic dataset ($n = 339$) of modern and recent historical hunter-gatherer populations (24) to extract calibration data to train the statistical models. This dataset is obviously geographically biased. The spatial distribution of ethnographically documented hunter-gatherers does not reflect the geographical area that is suitable for hunter-gatherers, because large areas previously occupied by foragers are dominated by agricultural populations from the Mid-Holocene onward. However, it has been shown that the ethnographic sample of hunter-gatherers is not biased in terms of their niche space (24). Therefore, these data are suitable for niche modeling including climate envelope modeling that are extrapolated to the geographical areas not recently occupied by hunter-gatherers.

For the calibration data, we excluded cases where subsistence is based on mutualistic relations with non–hunter-gatherers (SUBPOP = X). Because the isotope studies of human bone collagen indicate that the Pleistocene hunter-gatherers obtained, at most, 30% of their dietary protein from aquatic resources (46, 47), we also excluded populations whose main livelihood comes from aquatic resources (SUBSP = 3). In addition, we excluded populations that used horses (SYSTATE3 = 1), because mounted hunter-gatherers are unknown in the European Paleolithic record. To keep the simulated population densities conservative, we excluded populations that either move into and out of a central location that is maintained for more than 1 y or are completely sedentary (GRPPAT = 2). These groups usually live under high population densities. The exclusion means that the simulation assumes that the Pleistocene human populations in Europe were residentially mobile, an assumption commonly held by archaeologists. For a comparison, we present in *SI Text* and Fig. S2 a more relaxed simulation based on the calibration data that includes also semi- and fully sedentary groups.

Altogether, the calibration data includes information on 127 hunter-gatherer populations. Because this dataset gives information only on environments where the hunter-gatherers have existed in recent historical times, we added 120 pseudo-absence data points to the climate space where terrestrially adapted hunter-gatherers have not recently existed (e.g., extremely cold and extremely hot and dry) to enhance the performance of the statistical models (Fig. S3 and Dataset S1). Pseudo-absence data have information on climatic conditions and the hunter-gatherer density for each point is zero. The climate data for these points were obtained from the WorldClim database (48). Addition of pseudo-absence data to presence-only data are a standard procedure in ecological modeling (49, 50).

We used potential evapotranspiration (PET), water balance (WAB), and mean temperature of the coldest month (MCM) as predictors of the density (DENSITY) and presence/absence (DENSITY > 0) of the human population. PET and MCM values are directly available from the ethnographic dataset. WAB values were calculated as the difference between annual precipitation and PET.

To model the distribution and density of the human population, we used two frameworks: one predicting the range (presence/absence) of the human population and the other predicting population density. The human population occurrence was modeled as a binary response variable and density as a continuous response variable. To take into account the fact that different modeling algorithms give diverse predictions, the following six alternative techniques were used to relate human presence/absence and density with the explanatory climatic variables: generalized linear modeling (GLM) (51), generalized additive modeling (GAM) (52), support vector machines (SVM) (53, 54), classification tree analysis (CTA) (55, 56), random forest (RF) (57, 58), and generalized boosting methods (GBM) (59, 60). All of the methods were implemented using R statistical software (61). A more detailed description of these techniques is given in the *SI Text*.

Predicted probabilities of occurrence were converted to presence/absence predictions using the threshold value maximizing the sum of sensitivity and specificity (62) (*SI Text*).

The ability of the models to predict human population occurrence and density was assessed using cross-validation (70% random sample for calibration and 30% for validation; 500 repeats). The predictive power of the binary models was determined by testing the accuracy of predictions made for the validation dataset by calculating the area under the curve of a receiver operating characteristic plot (AUC) and the true skill statistic (TSS) (63). For density models, mean R^2 values were calculated. Predictive accuracies of the six models based on three climate variables are summarized in Table S1.

To further evaluate the ability of climate envelope modeling approach to correctly simulate hunter-gatherer populations, we simulated Australian hunter-gatherer population at 0.5 ky ago and compared the result to the historical, ethnographic, and archaeological estimates of population size at the European contact (64–66). This simulation is presented in the supplement (*SI Text* and Fig. S4).

Climate Model. The monthly average temperature and annual precipitation values for Europe were generated using a full last glacial cycle simulation (126 ky ago until the present day) with the CLIMBER-2-SICOPOLIS model system (27) that simulates climate at a temporal resolution of 1,000 y. Climate data were downscaled here to the resolution of 1.5° (longitude) × 0.75° (latitude) for a time slice of 30–13 ky ago using a GAM (52). The GAM used here was calibrated (28) using observations of the recent past climate (67, 68) and a short time slice simulation of the LGM (about 22 ky ago) using a relatively high-resolution general circulation model (CCSM4) (36). See *SI Text* for details. The temperature data at the spatial resolution of 1.5° × 0.75° were regridded to 0.375° × 0.250°. During the regridding process, monthly temperature values were lapsed by the pseudo adiabatic lapse rate (6.4 °C/km) to account for differences in average elevation between the fine-scale and coarse-scale grids (69) (Dataset S2). The problem with the climate model is that it cannot trace high-frequency climate variations. Therefore, for example, some of the cold events, such as Heinrich 1, do not show up in the model data.

Human Population Range and Density Model. The range of the human population for every 1,000 y between 30 and 13 ky ago was simulated by predicting presence/absence of humans for every 0.375° × 0.25° cell containing land area. This simulation was done by using the above-mentioned calibration model algorithms and climate predictor values derived from the climate simulation. The climate simulation based monthly average temperature, and annual precipitation values were used to calculate PET and WAB values. WAB was calculated as the difference between precipitation and potential evapotranspiration. PET was calculated as (70, 71)

$$\mathrm{PET} = 58.93 \times \mathrm{T}_{\mathrm{above\ 0\ °C}}$$

The results of different model algorithms were averaged by using ensemble averaging methods that have been shown to remarkably increase the robustness of forecasts (72). For binary models, majority vote was used. Majority vote is an ensemble forecasting method that assigns a presence prediction only when more than half of the models (i.e., >3) predicts a presence (73).

Next, population density was predicted for every 0.375° × 0.25° cell inside the modeled range. For density models, to average the results based on different algorithms, their median was calculated (consensus method) (72) for each cell.

To calculate the human population size in Europe every 1,000 y, we first calculated the land area of each cell. Here we took into account the systematic areal change of the 0.375° × 0.25° cells and the actual percentage of the land area in each cell. Next, we multiplied the predicted population density of the cell by the land area of the cell and summed these values to get the total population size.

To evaluate the uncertainty of population size estimates, we repeated the whole process from calibration model fitting to calculation of population size 500 times using each time a random sample (70%) of the training/calibration data. This procedure allowed us to calculate confidence limits for the simulated population size estimates. The set of modeling techniques and climate data were held constant throughout the process.

The changes in the percentage of inhabited land area in Europe between 30 and 13 ky ago were calculated by relating the summed land area of the inhabited cells to the total land area of the cells containing land (ice sheet included). To estimate the percentage of time the cell has been inhabited between 30 and 13 ky ago, we counted the number of 1,000-y intervals when the cell was inhabited (maximum 18) and related this to the total number of time intervals.

The ice sheets shown in Fig. 2 were drawn according to four sources (74–77). There is some overlap between the reconstructed ice sheets and the modeled population range, especially in the British Isles at 27 ky ago. This overlap may partly be due to the generalizing effect of using 1,000-y time intervals in climate and human population simulations and in the ice sheet reconstructions but may also reflect some inaccuracies in the modeled human populations ranges and/or ice sheet reconstructions.

We have taken into account eustatic changes in the sea level and the consequent changes in the land area of Europe by adjusting the sea level according to a global sea level change curve (78).

Archaeological Human Population Proxy. Previous approaches of prehistoric human distribution and niche modeling have trained the predictive models using archaeological site distribution data (79–81). By keeping our calibration model independent from the archaeological data, we are able to test our simulation with the archaeological data.

To evaluate the simulated human population range and density, they were compared with the archaeological population proxy. The population proxy is based on ^{14}C dates, and the dates extracted from the International Union for Quaternary Science (INQUA) Radiocarbon Paleolithic Europe Database v12 form the backbone of data (30). We also included several dates from other recently published sources. The reasoning behind such a dates-as-data approach is that reliable archaeological radiocarbon dates indicate human presence in the area and that the temporal variation in the frequencies of ^{14}C dates reflects changes in prehistoric population size (8–10, 31, 32, 82).

The dataset was critically evaluated using the information given in the INQUA database. We excluded (*i*) all dates that were qualified as unreliable or contaminated, (*ii*) dates without coordinates or laboratory reference, (*iii*) duplicate dates, (*iv*) dates with SEs greater than 5% of the mean ^{14}C age, (*v*) dates from gyttja, humus, peat, soil or soil organics, organic sediment, humic acid fraction of the sediment, and fossil timber, (*vi*) dates of marine origin, such as shell, marine shell, and molluscs, (*vii*) dates without a clear link to human activity, such as terminus ante and post quem, surface, above, up from, top, below and beneath of the cultural layer(s), minimum or maximum age of the layer, and beyond site, and (*viii*) dates of cave bear (*Ursus spelaeus*). In some cases, coordinates or even ages were corrected according to the original publication of the date. After the cleaning, the dataset contains 3,718 ^{14}C dates from 895 sites (Dataset S3).

The dates were calibrated using the OxCal 4.2 calibration program (83) and IntCal13 calibration curve (84). In the analyses, we used the calibrated median dates. For comparisons between the model and archaeological data, median dates were grouped in intervals of 1,000 y so that the modeled human range at 30 ky was compared with the spatial distribution of dates between 30,499 and 29,500 cal BP, the modeled range at 29,000 cal BP to the distribution of dates between 29,499 and 28,500 cal BP, and so forth.

Surovell et al. (31) argued that the younger findings are overrepresented relative to older findings in the archaeological record due to the time-dependent influence of destructive processes such as erosion and weathering. Similar time-dependent loss processes seem to affect geological and palaeontological data, as well as historical coin records (85, 86). Therefore, the temporal frequency distributions should be corrected for this taphonomic bias. Surovell et al. (31) proposed a model of taphonomic bias and suggested how to use it to correct the temporal frequency distributions. This method was evaluated, modified, and implemented in several subsequent studies (1, 32, 82, 86). Here, we used a taphonomic bias model modified by Williams (32). The temporal distribution of the taphonomically corrected number of dates was used as a proxy for relative changes in human population size between 30 and 13 ky ago, and this distribution was compared with the temporal distribution of modeled population sizes. See Fig. S5 for the comparison between raw temporal frequency distribution and the taphonomically corrected temporal frequency distribution of archaeological dates.

ACKNOWLEDGMENTS. We thank A. Lister, M. Fortelius, T. Rankama, H. Renssen, and F. Riede for discussions and comments on an earlier version of the manuscript; A. Ganopolski for providing the climate simulations of CLIMBER-2-SICOPOLIS; and W. Perttola for technical help with spatial analyses. M.T. acknowledges financial support from the Kone Foundation.

1. Lorenzen ED, et al. (2011) Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* 479(7373):359–364.
2. Mellars P, French JC (2011) Tenfold population increase in Western Europe at the Neandertal-to-modern human transition. *Science* 333(6042):623–627.
3. Derex M, Beugin M-P, Godelle B, Raymond M (2013) Experimental evidence for the influence of group size on cultural complexity. *Nature* 503(7476):389–391.
4. Powell A, Shennan S, Thomas MG (2009) Late Pleistocene demography and the appearance of modern human behavior. *Science* 324(5932):1298–1301.
5. Bromham L, Hua X, Fitzpatrick TG, Greenhill SJ (2015) Rate of language evolution is affected by population size. *Proc Natl Acad Sci USA* 112(7):2097–2102.
6. Henrich J (2004) Demography and cultural evolution: How adaptive cultural processes can produce maladaptive losses: The Tasmanian case. *Am Antiq* 69(2):197–214.
7. Frankham R (1996) Relationship of genetic variation to population size in wildlife. *Conserv Biol* 10(6):1500–1508.
8. Gamble C, Davies W, Pettitt P, Hazelwood L, Richards M (2005) The archaeological and genetic foundations of the European population during the Late Glacial: Implications for "agricultural thinking." *Camb Archaeol J* 15(02):193–223.
9. Steele J (2010) Radiocarbon dates as data: Quantitative strategies for estimating colonization front speeds and event densities. *J Archaeol Sci* 37(8):2017–2030.
10. Shennan S, et al. (2013) Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nat Commun* 4:2486.
11. Atkinson QD, Gray RD, Drummond AJ (2008) mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Mol Biol Evol* 25(2):468–474.
12. Li H, Durbin R (2011) Inference of human population history from individual whole-genome sequences. *Nature* 475(7357):493–496.
13. Schiffels S, Durbin R (2014) Inferring human population size and separation history from multiple genome sequences. *Nat Genet* 46(8):919–925.
14. Scally A, Durbin R (2012) Revising the human mutation rate: Implications for understanding human evolution. *Nat Rev Genet* 13(10):745–753.
15. Fu Q, et al. (2013) A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol* 23(7):553–559.
16. Hawks J (2008) From genes to numbers: Effective population sizes in human evolution. *Recent Advances in Palaeodemography*, ed Bocquet-Appel J-P (Springer, Dordrecht, The Netherlands), pp 9–30.
17. Kelly RL (2013) *The Lifeways of Hunter-Gatherers: The Foraging Spectrum* (Cambridge Univ Press, Cambridge, UK), 2nd Ed.
18. Binford LR (1980) Willow smoke and dogs' tails: Hunter-gatherer settlement systems and archaeological site formation. *Am Antiq* 45(1):4–20.
19. Birdsell JB (1953) Some environmental and cultural factors influencing the structuring of Australian aboriginal populations. *Am Nat* 87(834):171–207.
20. Hamilton MJ, Burger O, Walker RS (2012) Human ecology. *Metabolic Ecology*, eds Sibly RM, Brown JH, Kodric-Brown A (John Wiley & Sons, Chichester, UK), pp 248–257.
21. Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: Are bioclimate envelope models useful? *Glob Ecol Biogeogr* 12(5):361–371.
22. Araújo MB, Peterson AT (2012) Uses and misuses of bioclimatic envelope modeling. *Ecology* 93(7):1527–1539.
23. Svenning J-C, Fløjgaard C, Marske KA, Nógues-Bravo D, Normand S (2011) Applications of species distribution modeling to paleobiology. *Quat Sci Rev* 30(21-22):2930–2947.
24. Binford LR (2001) *Constructing Frames of Reference: An Analytical Method for Archaeological Theory Building Using Ethnographic and Environmental Data Sets* (Univ of California Press, Berkeley).
25. Currie DJ (1991) Energy and large-scale patterns of animal- and plant-species richness. *Am Nat* 137(1):27–49.
26. Franklin J (2010) *Mapping Species Distributions: Spatial Inference and Prediction* (Cambridge Univ Press, Cambridge, UK).
27. Ganopolski A, Calov R, Claussen M (2010) Simulation of the last glacial cycle with a coupled climate ice-sheet model of intermediate complexity. *Clim Past* 6(2):229–244.
28. Korhonen N, Venäläinen A, Seppä H, Järvinen H (2014) Statistical downscaling of a climate simulation of the last glacial cycle: Temperature and precipitation over Northern Europe. *Clim Past* 10(4):1489–1500.
29. Clark PU, et al. (2009) The Last Glacial Maximum. *Science* 325(5941):710–714.
30. Vermeersch PM (2005) European population changes during Marine Isotope Stages 2 and 3. *Quat Int* 137(1):77–85.

31. Surovell TA, Byrd Finley J, Smith GM, Brantingham PJ, Kelly R (2009) Correcting temporal frequency distributions for taphonomic bias. *J Archaeol Sci* 36(8):1715–1724.
32. Williams AN (2012) The use of summed radiocarbon probability distributions in archaeology: A review of methods. *J Archaeol Sci* 39(3):578–589.
33. Terberger T, Street M (2002) Hiatus or continuity? New results for the question of pleniglacial settlement in Central Europe. *Antiquity* 76(293):691.
34. Slimak L, et al. (2011) Late Mousterian persistence near the Arctic Circle. *Science* 332(6031):841–845.
35. Zwyns N, Roebroeks W, McPherron SP, Jagich A, Hublin J-J (2012) Comment on "Late Mousterian persistence near the Arctic Circle". *Science* 335(6065):167.
36. Gent PR, et al. (2011) The community climate system model, version 4. *J Clim* 24(19): 4973–4991.
37. Giorgetta MA, et al. (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *J Adv Model Earth Syst* 5(3):572–597.
38. Watanabe S, et al. (2011) MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geosci Model Dev* 4(4):845–872.
39. Vanmaercke M, Poesen J, Verstraeten G, de Vente J, Ocakoglu F (2011) Sediment yield in Europe: Spatial patterns and scale dependency. *Geomorphology* 130(3-4):142–161.
40. Poesen JWA, Hooke JM (1997) Erosion, flooding and channel management in Mediterranean environments of southern Europe. *Prog Phys Geogr* 21(2):157–199.
41. Jennings R, Finlayson C, Fa D, Finlayson G (2011) Southern Iberia as a refuge for the last Neanderthal populations. *J Biogeogr* 38(10):1873–1885.
42. Roche J (1964) Le Paléolithique supérieur portugais. Bilan de nos connaissances et problèmes. *Bull Société Préhistorique Fr Études Trav* 61(1):11–27.
43. Aubry T, Llach XM, Sampaio JD, Sellami F (2002) Open-air rock-art, territories and modes of exploitation during the Upper Palaeolithic in the Côa Valley (Portugal). *Antiquity* 76(291):62–76.
44. Aubry T, Luís L, Mangado Llach J, Matias H (2012) We will be known by the tracks we leave behind: Exotic lithic raw materials, mobility and social networking among the Côa Valley foragers (Portugal). *J Anthropol Archaeol* 31(4):528–550.
45. Bocquet-Appel J-P, Demars P-Y, Noiret L, Dobrowsky D (2005) Estimates of Upper Palaeolithic meta-population size in Europe from archaeological data. *J Archaeol Sci* 32(11):1656–1668.
46. Richards MP, Pettitt PB, Stiner MC, Trinkaus E (2001) Stable isotope evidence for increasing dietary breadth in the European mid-Upper Paleolithic. *Proc Natl Acad Sci USA* 98(11):6528–6532.
47. Drucker D, Bocherens H (2004) Carbon and nitrogen stable isotopes as tracers of change in diet breadth during Middle and Upper Palaeolithic in Europe. *Int J Osteoarchaeol* 14(3-4):162–177.
48. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25(15):1965–1978.
49. Phillips SJ, et al. (2009) Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. *Ecol Appl* 19(1):181–197.
50. Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: How, where and how many? *Methods Ecol Evol* 3(2):327–338.
51. McCullagh P, Nelder JA (1989) *Generalized Linear Models* (Chapman & Hall/CRC, Boca Raton, FL).
52. Hastie T, Tibshirani R (1990) *Generalized Additive Models* (Chapman & Hall/CRC, Boca Raton, FL).
53. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297.
54. Drucker H, Burges CJC, Kaufman L, Smola AJ, Vapnik V (1997) Support vector regression machines. *Advances in Neural Information Processing Systems 9*, eds Mozer MC, Jordan MI, Petsche T (MIT Press, Cambridge, MA), pp 155–161.
55. Breiman L, Friedman J, Stone CJ, Olshen RA (1984) *Classification and Regression Trees* (Chapman and Hall/CRC, New York), 1st Ed.
56. Venables WN, Ripley BD (2010) *Modern Applied Statistics with S* (Springer, New York), 4th Ed.
57. Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32.
58. Prasad A, Iverson L, Liaw A (2006) Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems (N Y)* 9(2):181–199.
59. Ridgeway G (1999) The state of boosting. *Comput Sci Stat* 31:172–181.
60. Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. *J Anim Ecol* 77(4):802–813.
61. R Development Core Team (2014) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna).
62. Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28(3):385–393.
63. Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS): Assessing the accuracy of distribution models. *J Appl Ecol* 43(6):1223–1232.
64. White JP, Mulvaney DJ (1987) How many people? *Australians to 1788*, eds Mulvaney DJ, White JP (Fairfax, Syme & Weldon Associates, Broadway, Australia), pp 114–117.
65. Mulvaney DJ, Kamminga J (1999) *Prehistory of Australia* (Allen & Unwin, St Leonards, Australia).
66. Williams AN (2013) A new population curve for prehistoric Australia. *Proc R Soc B Biol Sci* 280(1761):20130486.
67. Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int J Climatol* 25(6):693–712.
68. Xie P, Arkin PA (1997) Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull Am Meteorol Soc* 78(11):2539–2558.
69. Vajda A, Venäläinen A (2003) The influence of natural conditions on the spatial variation of climate in Lapland, northern Finland. *Int J Climatol* 23(9):1011–1022.
70. Skov F, Svenning J-C (2004) Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography* 27(3):366–380.
71. Holdridge LR (1967) *Life Zone Ecology* (Tropical Science Center, San Jose, Costa Rica).
72. Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends Ecol Evol* 22(1):42–47.
73. Gallien L, Douzet R, Pratte S, Zimmermann NE, Thuiller W (2012) Invasive species distribution models – how violating the equilibrium assumption can create new insights. *Glob Ecol Biogeogr* 21(11):1126–1136.
74. Boulton G, Dongelmans P, Punkari M, Broadgate M (2001) Palaeoglaciology of an ice sheet through a glacial cycle: The European ice sheet through the Weichselian. *Quat Sci Rev* 20(4):591–625.
75. Clark CD, Hughes ALC, Greenwood SL, Jordan C, Sejrup HP (2012) Pattern and timing of retreat of the last British-Irish Ice Sheet. *Quat Sci Rev* 44:112–146.
76. Lambeck K, Purcell A, Zhao J, Svensson N-O (2010) The Scandinavian Ice Sheet: From MIS 4 to the end of the Last Glacial Maximum. *Boreas* 39(2):410–435.
77. Svendsen JI, et al. (2004) Late Quaternary ice sheet history of northern Eurasia. *Quat Sci Rev* 23(11–13):1229–1271.
78. Peltier WR, Fairbanks RG (2006) Global glacial ice volume and Last Glacial Maximum duration from an extended Barbados sea level record. *Quat Sci Rev* 25(23–24):3322–3337.
79. Banks WE, et al. (2008) Human ecological niches and ranges during the LGM in Europe derived from an application of eco-cultural niche modeling. *J Archaeol Sci* 35(2):481–491.
80. Banks WE, et al. (2011) Eco-cultural niches of the Badegoulian: Unraveling links between cultural adaptation and ecology during the Last Glacial Maximum in France. *J Anthropol Archaeol* 30(3):359–374.
81. Beeton TA, Glantz MM, Trainer AK, Temirbekov SS, Reich RM (2014) The fundamental hominin niche in late Pleistocene Central Asia: A preliminary refugium model. *J Biogeogr* 41(1):95–110.
82. Kelly RL, Surovell TA, Shuman BN, Smith GM (2013) A continuous climatic impact on Holocene human population in the Rocky Mountains. *Proc Natl Acad Sci USA* 110(2):443–447.
83. Bronk Ramsey C (2009) Bayesian analysis of radiocarbon dates. *Radiocarbon* 51(1): 337–360.
84. Reimer PJ, et al. (2013) IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* 55(4):1869–1887.
85. Surovell TA, Brantingham PJ (2007) A note on the use of temporal frequency distributions in studies of prehistoric demography. *J Archaeol Sci* 34(11):1868–1877.
86. Peros MC, Munoz SE, Gajewski K, Viau AE (2010) Prehistoric demography of North America inferred from radiocarbon data. *J Archaeol Sci* 37(3):656–664.
87. Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc Series B Stat Methodol* 73(1):3–36.
88. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F (2014) e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R package version 1.6-4. Available at CRAN.R-project.org/package=e1071. Accessed May, 2014.
89. Therneau T, Atkinson B, Ripley B (2014) rpart: Recursive partitioning and regression trees, R package version 4.1-8. Available at CRAN.R-project.org/package=rpart. Accessed May, 2014.
90. Liaw A, Wiener M (2002) Classification and regression by randomForest. *R News* 2(3): 18–22.
91. Friedman JH (2001) Greedy function approximation: A gradient boosting machine. *Ann Stat* 29(5):1189–1232.
92. VanDerWal J, Falconi L, Januchowski S, Shoo L, Storlie C (2014) SDMTools: Species Distribution Modelling Tools: Tools for processing data associated with species distribution modelling exercises. R package version 1.1-221. Available at CRAN.R-project.org/package=SDMTools. Accessed May, 2014.
93. Petoukhov V, et al. (2000) CLIMBER-2: A climate system model of intermediate complexity. Part I: Model description and performance for present climate. *Clim Dyn* 16(1):1–17.
94. Petoukhov V, et al. (2005) EMIC Intercomparison Project (EMIP–CO2): Comparative analysis of EMIC simulations of climate, and of equilibrium and transient responses to atmospheric CO2 doubling. *Clim Dyn* 25(4):363–385.
95. Greve R (1997) Application of a polythermal three-dimensional ice sheet model to the Greenland ice sheet: Response to steady-state and transient climate scenarios. *J Clim* 10(5):901–918.
96. Calov R, Ganopolski A, Claussen M, Petoukhov V, Greve R (2005) Transient simulation of the last glacial inception. Part I: Glacial inception as a bifurcation in the climate system. *Clim Dyn* 24(6):545–561.
97. Berger A (1978) Long-term variations of caloric insolation resulting from the earth's orbital elements. *Quat Res* 9(2):139–167.
98. Schneider von Deimling T, Held H, Ganopolski A, Rahmstorf S (2006) Climate sensitivity estimated from ensemble simulations of glacial climate. *Clim Dyn* 27(2-3): 149–163.
99. Oerlemans J (2005) Extracting a climate signal from 169 glacier records. *Science* 308(5722):675–677.
100. Intergovernmental Panel on Climate Change, ed. (2014) Information from paleoclimate archives. *Climate Change 2013: The Physical Science Basis* (Cambridge Univ Press, Cambridge, UK), pp 383–464.
101. Iakovleva L, Djindjian F, Maschenko EN, Konik S, Moigne A-M (2012) The late Upper Palaeolithic site of Gontsy (Ukraine): A reference for the reconstruction of the hunter–gatherer system based on a mammoth economy. *Quat Int* 255:86–93.
102. Starkovich BM, Stiner MC (2010) Upper Palaeolithic animal Exploitation at Klissoura Cave 1 in Southern Greece: Dietary Trends and Mammal Taphonomy. *Eurasian Prehistory* 7(2):107–132.
103. Starkovich BM (2012) Fallow deer (Dama dama) hunting during the Late Pleistocene at Klissoura Cave 1 (Peloponnese, Greece). *Mitteilungen Ges Für Urgesch* 21:11–36.

Supporting Information

Tallavaara et al. 10.1073/pnas.1503784112

SI Text

Calibration Modeling Techniques.

i) GLMs are mathematical extensions of linear models (51), which can handle nonlinear relationships and different types of statistical distributions characterizing spatial data and are technically closely related to traditional practices used in linear modeling and ANOVA. For the density model, linear and second- and third-order polynomial terms and, for the binary model, linear and second-order polynomial terms were computed with family specification = Poisson to provide the human population density and probability of human population occurrence in each grid square, as a response to the three climatic variables.

ii) GAMs are nonparametric extensions of GLMs (52). GAMs provide a flexible data-driven class of models based on a cubic-spline smoother with 4 df that permit both linear and complex additive response shapes, as well as a combination of the two within the same model. The smooth functions are computed independently for each explanatory variable and added to construct the final model. GAM was implemented using the package mgcv (87) in R (61). The following smoothing parameters were used for the density model: k = 3 (MCM), k = 4 (PET), k = 4 (WAB). The following smoothing parameters were used for the binary model: k = 3 (MCM), k = 3 (PET), k = 3 (WAB). Family specification was set to Poisson.

iii) SVM is a machine learning method able to deal with complex nonlinear relations between variables. The SVM constructs a hyperplane or set of hyperplanes in a high-dimensional space, which can be used for classification and regression (53, 54). The SVM transforms a nonlinear relationship to linear through its ability to map the problem into a higher dimension space using a kernel trick. The SVM was implemented using package e1071 (88). A radial basis kernel was used in both binary (C-classification) and density (eps-regression) models. Optimal values for gamma and cos parameters were searched using cross-validation using tune function. For binary model, values are gamma = 2 and cost = 8 and for the density model are gamma = 1 and cost = 1.

iv) CTA is an alternative to regression techniques. It is based on classification trees (55) and uses recursive partitioning to split the data into progressively smaller, homogenous subsets until a termination is reached (56). The optimal length of the tree is selected by a procedure of 10 cross-validations. The advantage of CTA is that it allows the capturing of nonadditive behavior and complex interactions. CTA was implemented using the package rpart (89). Optimal values for the minimum number of observations that must exist in a node for a split to be attempted (minsplit) were searched using cross-validation. For the density model, the minsplit value was set to 20 and for the binary model was set to 50. Method was set to Poisson.

v) RF is a machine learning method (57). RF generates hundreds of random trees. A selective algorithm limits the number of implemented parameters in each tree. A training set for each tree is chosen as many times as there are observations, among the whole set of observations. After the trees have been built, data are entered into them and each grid square is classified by all of the trees. At the end of the run, the classification given by each tree is considered as a "vote," and the classification of a grid square corresponds to the majority vote among all trees (58). RF was implemented using the package randomForest v. 4.6–10 (90).

vi) GBM is a machine learning method that estimates the relationship between a response variable and its predictors without a priori specification of a data model (59). Boosted regression trees (60), the GBM technique used here, is an advanced form of regression that automatically incorporates interactions between predictors and is capable of modeling complex nonlinear functions. Boosting is a numerical optimization technique for minimizing a loss function (such as deviance) by adding at each step a new tree that best reduces the loss function (59, 60). Here, GBM was implemented using the package gmb (91). The distribution parameter was set to Poisson. The total number of trees was set to 1,000, number of cross-validation folds to 4, shrinkage parameter to 0.005, and the maximum depth of variable interaction to 3. In addition, gbm.perf function was used to estimates the optimal number of boosting iterations.

Optimal Threshold Calculation. Binary models were initially set to predict probabilities of occurrence. If the probability was higher than a certain threshold, the observation was classified as present (1) and otherwise as absent (0). A model-specific optimal threshold was calculated using calibration data so that the threshold would maximize the sum of the percentage of correctly identified occurrences (sensitivity) and absences (specificity). Calculation was performed using the optim.thresh function in the SDMTools package (92). This function often returns a range in thresholds that are equal for a given data and model. In such cases, the 95th percentile was selected as the threshold value, which means that our simulations of the human population range are relatively conservative.

Climate Model. The climate predictors for Europe were generated using a full last glacial cycle simulation (126 ky ago until the present day) with the CLIMBER-2-SICOPOLIS model system (27). CLIMBER-2 is an Earth system model of intermediate complexity (EMIC). It includes model components for the atmosphere, ocean, sea ice, land surface, terrestrial vegetation, and ice sheets (93). The reason for using an EMIC instead of an Earth system model of full complexity (ESM) is that the time period considered is too long for ESM simulations with adequate spatial resolution. Therefore, downscaling the simulation results would have been necessary regardless of the choice of the simulation model. The resolution of the simulation data (10° in latitude and 51° in longitude in the atmospheric component) is very low but nevertheless sufficient for simulating the seasonal variability of many climate variables (93) and the climate response to changes in different types of forcing and boundary conditions within the range of ESM simulations (94). Of particular interest for us is the ability of this EMIC to simulate observed ice sheet evolution to obtain a more realistic climatology of temperature and precipitation in regions adjacent to land-ice masses. The CLIMBER-2 model has been coupled with a high-resolution 3D thermomechanical ice sheet model (SICOPOLIS) (95). The resolution of the ice sheet model is 1.5° × 0.75°, and the model domain covers the Northern Hemisphere from 21°N to 85.5°N. The atmosphere and ice sheets are coupled using a physically based energy and mass balance interface (SEMI) (96). The simulations were forced by the Earth's orbital parameter variations (97), and atmospheric greenhouse gas concentrations were derived from the Vostok ice core. The modulating effect of atmospheric dust on radiative forcing was parameterized in proportion to the simulated global ice volume (98).

The annual mean temperature and total precipitation over Europe as simulated by CLIMBER-2 were downscaled here to

resolution of 1.5° (longitude) by 0.75° (latitude) for a time slice of 30–13 ky ago using a GAM (52). The GAM used here was calibrated (28) using observations of the recent past climate (67, 68) and a short time-slice simulation of the LGM (about 21 ky ago) using a relatively high-resolution general circulation model (CCSM4) (36). The explanatory variables in the GAM were temperature and precipitation of CLIMBER-2, and elevation, distance to ice sheet, slope direction, and slope angle of the SICOPOLIS ice sheet model.

The mean temperature of the coldest month $T_C(x,y,t)$ over Europe in each grid point (x,y) during the period 30–13 ky ago (t) with a time step of 1,000 y was computed as

$$T_C(x,y,t) = T_A(x,y,t) + \left[T_{Cp}(x,y) - T_{Ap}(x,y)\right] \times \left[1 - \frac{|T_{EAp} - T_{EA}(t)|}{T_{EAp} - T_{EAL}}\right] + \left[T_{CL}(x,y) - T_{AL}(x,y)\right] \times \left[\frac{|T_{EAp} - T_{EA}(t)|}{T_{EAp} - T_{EAL}}\right],$$

where $T_A(x,y,t)$ is the downscaled annual mean temperature, $T_{Cp}(x,y) - TAp(x,y)$ is the difference between the mean temperature of the coldest month and the annual mean temperature in each grid in the observed recent past (67, 68), and $T_{CL}(x,y) - TAL(x,y)$ is the corresponding difference in the general circulation model simulation of the LGM (36). T_{EAL}, $T_{EA}(t)$, and T_{EAp} are the downscaled annual mean temperatures over Europe during the LGM, each millennium from 30 to 13 ky ago, and observed past, respectively.

Simulation of the Australian Hunter-Gatherer Population at 0.5 ky Ago. To further evaluate the ability of the climate envelope modeling approach to simulate hunter-gatherer population densities, we simulated Australian population before the European contact at around 0.5 ky ago. The most recent estimates of Australian population size suggest that at the time of European contact (AD 1788), the hunter-gatherer population size was between 750,000 and 1.1 million and slightly higher at 0.5 ky ago (64–66). To simulate the Australian hunter-gatherer population, we extracted modern climate data from WorldClim database (48). As the Australian temperatures at 0.5 ky ago are estimated to be 0.5 °C lower than the current temperatures (99, 100), we subtracted 0.5 from all of the modern temperature values used in the simulation. Approximately similar climatic conditions prevailed also during the European contact in the 18th century, as the temperatures started to rise during the 19th and 20th century (99, 100).

We performed the Australian simulations using two different calibration sets. The first includes only terrestrially adapted groups. It also includes GRPPAT = 2; type hunter-gatherers as these are known in the Australian ethnographic record (24). Because the Australian ethnographic and historical record includes also aquatically adapted hunter-gatherers (24), whose population densities can be higher than the densities of terrestrially adapted populations living in similar climate, the simulation based on this kind of calibration data would underestimate Australian population size. Therefore, we performed the simulation also with the calibration data that also includes aquatically adapted hunter-gatherers (SUBSP = 3).

The results of these simulations are shown in Fig. S4. The simulation based on more realistic calibration data that also include aquatic foragers gives a precontact population size estimate of 780,000. This is within the suggested range at the time of the European contact, but slightly lower than the suggested population size of 1.2 million at 0.5 ky ago (66). The simulation suggests that areas in western and central Australia would have been uninhabited mainly due to the extremely negative water balance (Fig. S4). However, it is known that humans lived also in these areas. The reason for the simulated pattern relates to our selection of pseudo-absence data, which include locations in Africa and Arabian Peninsula, where the climate conditions resemble those of central Australia. Due to these pseudo-absence data points, the binary distribution model (humans present/absent) predicts uninhabited areas in Australia. When the selection of the optimal threshold value (see above) is relaxed from the 95th to the 5th percentile of the range of the optimal threshold values, population size increases to 800,000, and the uninhabited area decreases but does not disappear.

Despite the small discrepancy between the observed and simulated presence of humans, the simulated population size falls well within the range of historical and ethnographic estimates at the time of contact (64, 65) and is close to the size estimates based on archaeological and genetic data (66). Also the pattern in simulated population densities appears as expected and corresponds quite well the distribution archaeological data in Australia (66).

Simulation of the European Population 30–13 ky Ago Using Larger Ethnographic Calibration Data. The European archaeological record may not totally exclude the possibility of semi- or fully sedentary populations. For example, substantial mammoth bone dwellings indicate that even in the Eastern European tundra-steppe zone, residential mobility was not very high (101) and in the more suitable region of the Mediterranean Greece, the characteristics of the bone assemblage of Klissoura Cave 1 allow for the possibility of year-round settlement during certain periods of the Paleolithic (102, 103). Therefore, we present here a simulation that is based on the ethnographic calibration data that also contain the semisedentary or sedentary groups (GRPPAT = 2). Such groups usually live under high population densities, and consequently, the simulation will yield higher population densities in the Pleistocene Europe than if these groups were excluded from the calibration data. Otherwise, the calibration data are the same as used in the primary simulation. One outlying GRPPAT = 2-type group, Clear Lake Pomo, was excluded from this calibration data because its density is exceptionally high. The results of the simulation are shown in Fig. S2.

This simulation suggests that the human population size in Europe was about 500,000 at 30 ky ago, 200,000 during its minimum at 23 ky ago, and almost 700,000 at 13 ky ago, during the Greenland interstadial 1. The mean population density in the inhabited area varied between 4.3 and 8.0 per 100 km^2.

As semisedentary or sedentary (GRPPAT = 2) groups are included in this calibration data, it is reasonable to ask to what extent the simulated population densities of the Pleistocene Europe would indicate at least a semisedentary mobility pattern. To illustrate this, we fitted a logistic regression model to the ethnographic data to explain the type of mobility pattern (mobile vs. semi- or full sedentary) with population density and PET that turned out to be the only statistically significant climate variable in this context.

Next, we used this model to hindcast occurrence probability of semi- or full sedentary mobility pattern over European land area at two time slices (23 and 13 ky ago). We used the simulated population density and the PET values from the downscaled CLIMBER-2 as predictor variables. The results of this simulation are shown in Fig. S2. They show that semi- or full sedentary mobility pattern would have occurred only in the climatically most suitable areas along the coasts.

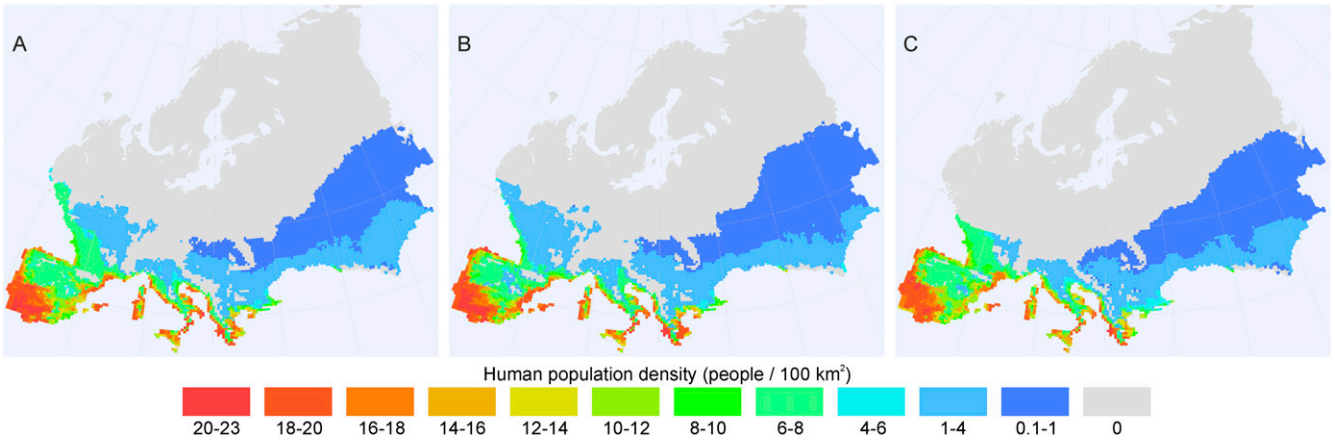

Fig. S1. Population simulations based on LGM snapshot (~21 ky) climate data downscaled from three different general circulation models. (*A*) Simulation based on the CCSM4 climate data (36). (*B*) Simulation based on the MPI-ESM climate data (37). (*C*) Simulation based on the MIROC-ESM climate data (38). Downscaled data were obtained from the WorldClim database (www.worldclim.org/). All three simulations show the pattern where the Iberian Peninsula and the Mediterranean region are climatically most suitable for hunter-gatherers.

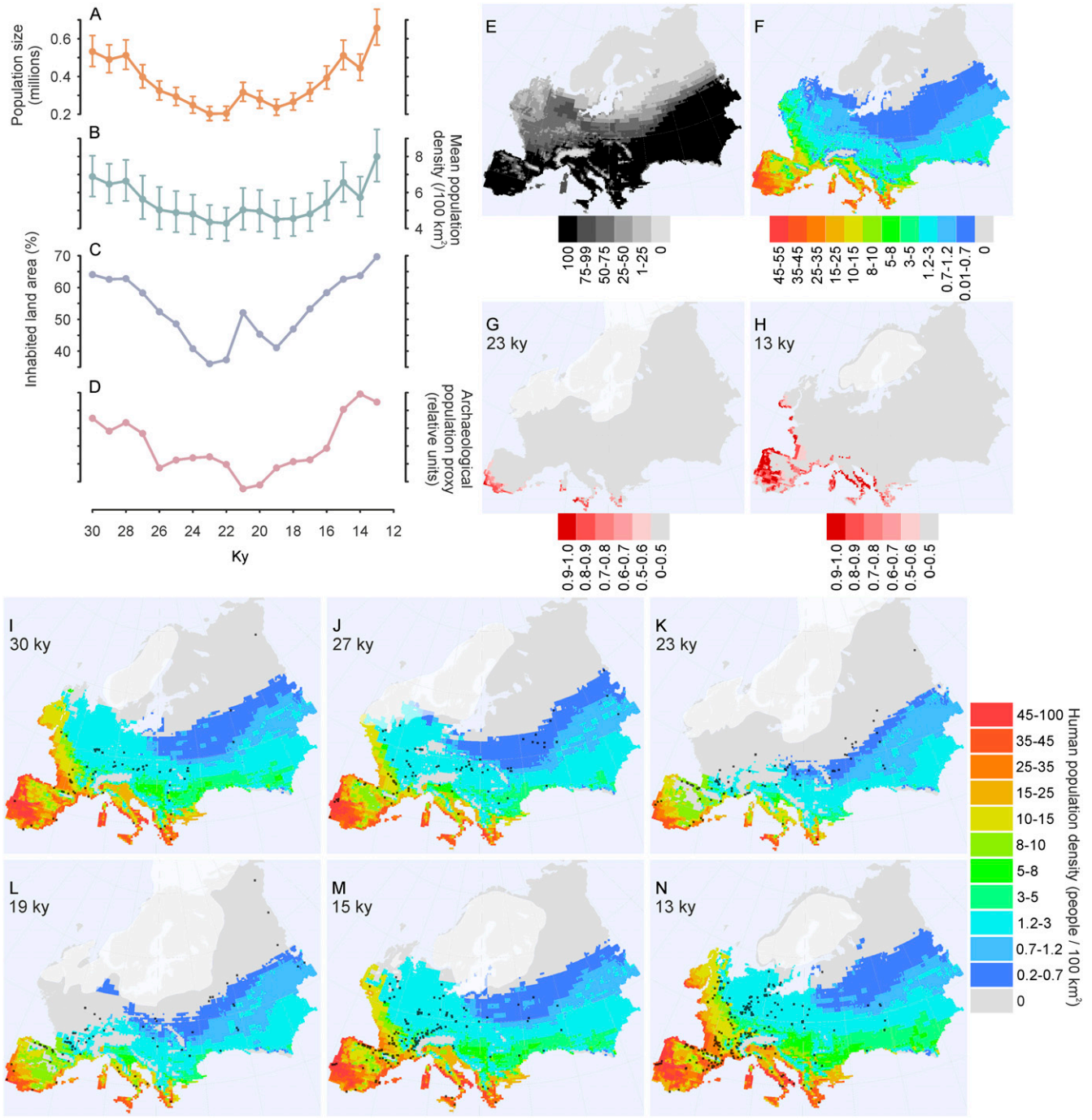

Fig. S2. Human population simulation in Europe based on the calibration data that includes also semi- and fully sedentary hunter-gatherer populations (GRPPAT = 2). (*A*) Simulated human population size in Europe. Error bars show the resampling-based confidence limits (95%). (*B*) Simulated mean density in the inhabited area of Europe. Error bars show the resampling-based confidence limits (95%). (*C*) Changes in the percentage of potentially inhabited land area in Europe. (*D*) Archaeological population size proxy based on the taphonomically corrected number of dates. (*E*) Percentage of time the area has potentially been inhabited between 30 and 13 ky ago. (*F*) Mean population density (people/100 km^2) between 30 and 13 ky ago. (*G* and *H*) Probability of semi- or fully sedentary mobility patterns at two time slices. (*I*–*N*) Simulated human population range and density compared with the spatial distribution of archaeological sites during six time intervals from 30 to 13 ky ago. Archaeological sites are indicated with black dots and in each time slice they represent sites dated within 1,000-y bins.

Fig. S3. The distribution of the calibration or training data.

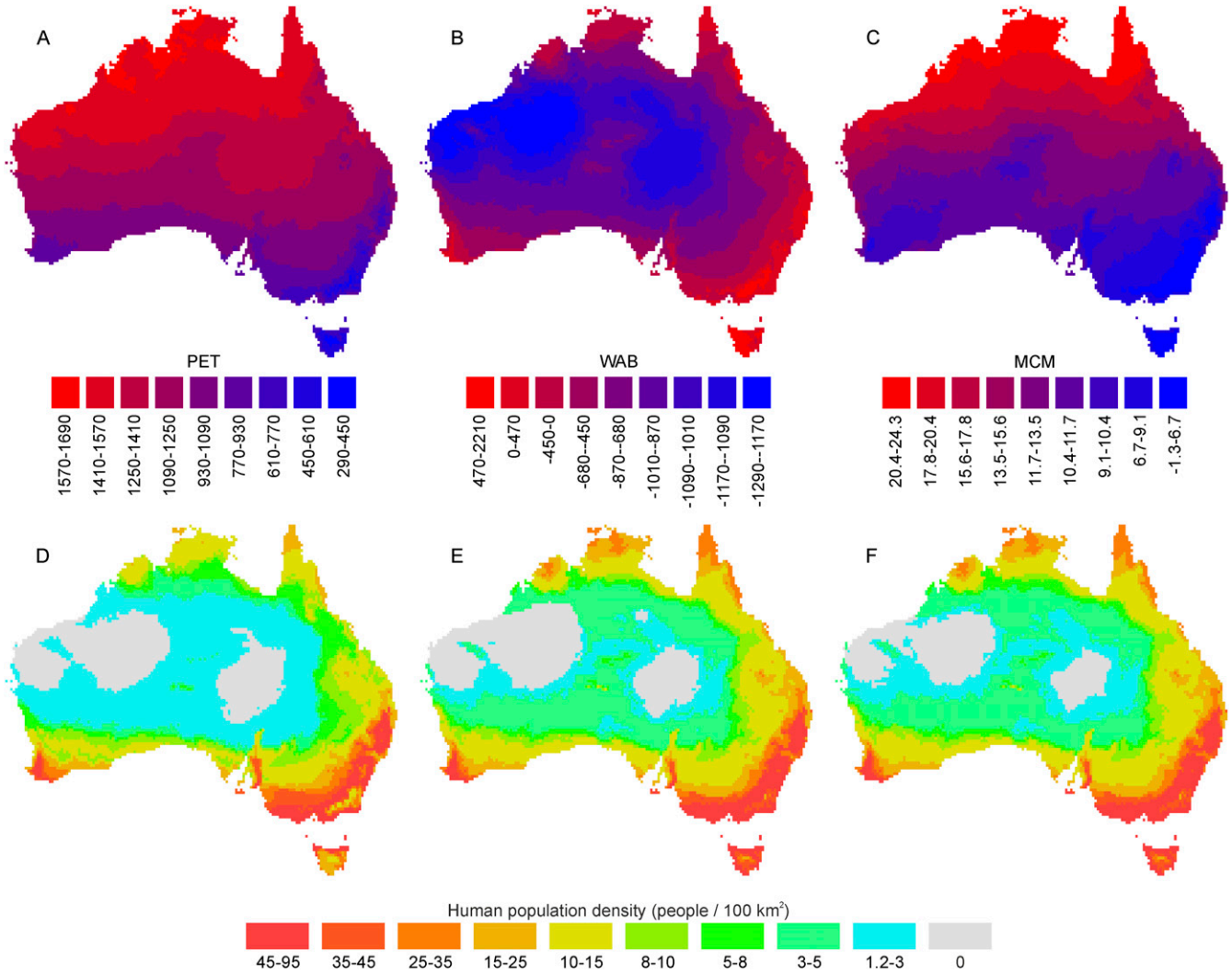

Fig. S4. Climate variability and population simulation in Australia at 0.5 ky ago. (*A*) Potential evapotranspiration. (*B*) Water balance. (*C*) Mean temperature of the coldest month. (*D*) Population simulation using only terrestrial hunter-gatherers (SUBSP = 1 and 2) as calibration data (population size = 620,000). (*E*) Population simulation using terrestrial and aquatic hunter-gatherers as calibration data (population size = 780,000). (*F*) The same as *E*, but based on more relaxed optimal threshold value (population size = 800,000).

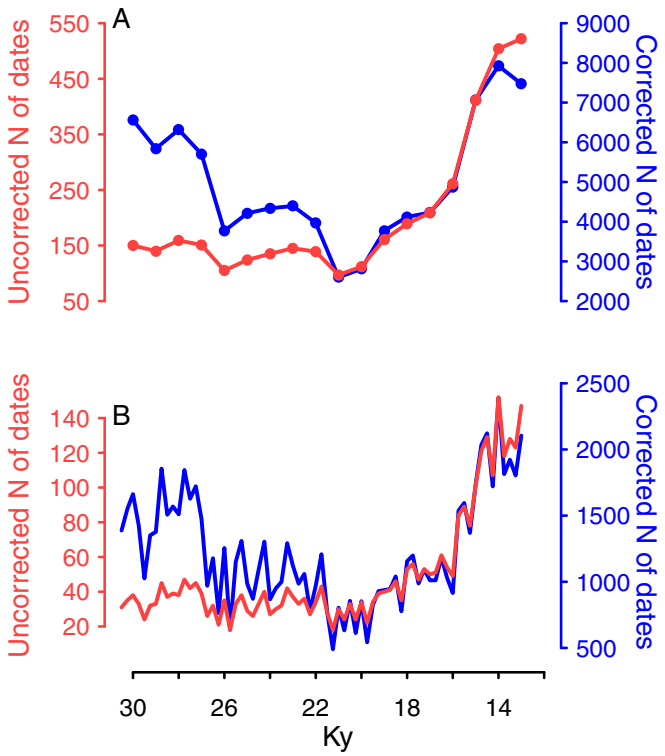

Fig. S5. Taphonomically uncorrected and corrected temporal distributions of archaeological dates with bins of 1,000 (*A*) and 250 (*B*) y.

Table S1. Predictive accuracies of the six human occurrence and density models based on three climate variables

Modeling technique	AUC	TSS	R^2
CTA	0.991	0.969	0.533
GAM	0.992	0.971	0.569
GBM	0.992	0.975	0.619
GLM	0.994	0.965	0.503
RF	0.996	0.962	0.614
SVM	0.995	0.954	0.543

Values are calculated from validation data. Calibration data were randomly split to training (70%) and validation (30%) sets. Models were built using training data, and their predictive abilities were assessed using new data (validation set). This was repeated 500 times, and the reported accuracies are mean values.

Other Supporting Information Files

Dataset S1 (XLS)
Dataset S2 (XLS)
Dataset S3 (XLS)

Paper V

Manuscript under revision for Atmospheric Chemistry and Physics

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

Adding value to Extended-range Forecasts in Northern Europe by Statistical Post-processing Using Stratospheric Observations

Natalia Korhonen[1,2], Otto Hyvärinen[1], Matti Kämäräinen[1], David S. Richardson[3], Heikki Järvinen[4], Hilppa Gregow[1]

[1]Weather and Climate Change Impact Research, Finnish Meteorological Institute, Helsinki, 00101, Finland
[2]Doctoral Programme in Atmospheric Sciences, University of Helsinki, Finland
[3]ECMWF, Reading, UK
[4]Institute for Atmospheric and Earth System Research/Physics, University of Helsinki, 00014, Finland

Correspondence to: Natalia Korhonen (Natalia.Korhonen@fmi.fi)

Abstract. The skill scores of the Extended-Range Forecasts (ERF) of the European Centre for Medium-Range Weather Forecasts (ECMWF) are still quite modest for the forecast weeks 3–6 in Northern Europe. As there are known stratospheric precursors impacting the surface weather with potential to improve ERFs, we aim to quantify the effect of these predictors and post-process the ERFs with them.

During boreal winter the quasi-biennial oscillation (QBO) affects the polar vortex; the easterly (westerly) QBO often coincides with weaker (stronger) than average polar vortex. Consequently, the weaker (stronger) than average stratospheric polar vortex is connected to negative (positive) Arctic Oscillation (AO) and colder (warmer) than average surface temperatures in Northern Europe. We developed a stratospheric wind indicator, *SWI*, based on the previous months' stratospheric wind observations and the phase of the AO during the following weeks. We demonstrate that there was a statistically significant difference in the observed surface temperature within the 3–6 weeks depending on the *SWI* at the start of the forecast. These temperature anomalies were underestimated by the ECMWF's reforecasts.

When our new *SWI* was applied in post-processing the ECMWF's two-week mean temperature reforecasts for weeks 3–4 and weeks 5–6 in Northern Europe during boreal winter, the skill scores of those weeks were slightly improved. This indicates there is some room to improve the ERFs, if the stratosphere-troposphere links were better captured in the modelling.

1 Introduction

Extended-range forecasts (ERF; lead time up to 46 days) by dynamical models have been developed since the 1990s with the aim to fill the gap between the medium-range weather forecasts and the seasonal forecasts. It is known that ERF skills are still rather modest in forecast weeks 3–6 especially in the Northern latitudes. If the skill of the forecasts improves, ERFs have the potential to become an essential element in climate services e.g., in the form of early warnings of climatic extremes. In an academic project CLIPS (CLImate services supporting Public mobility and Safety), climatic impact outlooks and early

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

warnings of extremes (CLIPS forecasts) were developed by employing the ERF datasets (Ervasti et al. 2018). The CLIPS forecasts were co-designed with the general public in Finland and experimented during a one year living lab. As many industries, e.g., energy and food production as well as users from the general public considered they could use and would benefit from reliable ERFs (Ervasti et al. 2018), development of more skillful ERFs is clearly needed.

The European Centre for Medium-Range Weather Forecasts (ECMWF) has produced ERFs routinely since March 2002 (Vitart 2014). The verification results of the ECMWF model's ERF (Buizza and Leutbecher 2015; Vitart 2014) on a sub-continental and a regional scale (e.g., Monhart et al. 2018) demonstrated predictive skill beyond 2 weeks for temperature reforecasts over Northern Europe. ECMWF uses bias correction of the mean in their automatic products, removing the mean bias computed from the reforecasts, depending on the time of the year (Buizza and Leutbecher 2015). We consider the bias over Northern Europe not to be dependent only on the time of the year but also on the prevailing weather pattern, and therefore, we aim to explore whether known teleconnections such as the strength of the stratospheric polar vortex, the phase of the Arctic Oscillation (AO), and the phase of the Quasi-Biennial Oscillation (QBO) could be used in improving the forecasts.

The stratospheric polar vortex is an upper-level low-pressure area that forms over both the northern and southern poles during winter due to the growing temperature gradient between the pole and the tropics. Strong westerly winds circulate the polar vortex, isolating the gradually cooling polar cap air. The strength of the northern polar vortex varies from year to year and can be indicated by, e.g., the zonal mean zonal wind (ZMZW) at 60 °N and 10 hPa or polar cap temperatures. The stronger the circumpolar winds and the colder the polar cap temperatures are, the stronger is the polar vortex. Planetary waves from the troposphere disturb the northern stratospheric polar vortex, leading to meandering and weakening of the westerlies and occasionally to reverse, i.e., easterly flow (Schoeberl, 1978). This weakening of the stratospheric polar vortex also leads to warming of the polar cap temperatures, sometimes even > 30–40 K within several days. A warming of this magnitude together with a reversal of the ZMZW at 10 hPa at 60 °N is commonly defined as a major sudden stratospheric warming (SSW), albeit other definitions also occur (Butler et al. 2015).

At the surface, the AO index, which is constructed from the daily 1000 hPa geopotential height field over 20°N–90°N against monthly mean values, is affected by the strength of the polar vortex with a time lag of about two to three weeks (Baldwin and Dunkerton 1999). A strong polar vortex is characterized by lower than average surface pressure in the Arctic, positive AO index, and strong westerly winds keeping the cold Arctic air locked in the polar region and bringing milder and wetter than average weather to Northern Europe (Limpasuvan et al. 2005). In contrast, a weak polar vortex is characterized by higher than average surface pressure in the Arctic, negative AO index, and the meandering and/or weakening of the polar jet stream and tropospheric jet stream enabling cold arctic/polar air outbreaks to Northern Europe (Thompson et al. 2002, Tomassini et al. 2012).

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

During boreal winters, the strength of the stratospheric polar vortex influences the surface weather in the Northern Hemisphere within weeks or months (Baldwin and Dunkerton 2001, Kidston et al. 2015), hence holding a potential for forecasting in that time scale. When making forecasts based on the strength of the polar vortex, a noteworthy phenomenon is also the QBO, a quasiperiodic oscillation of the equatorial zonal wind between downwards propagating easterlies and westerlies in the tropical stratosphere with a mean period of 28 to 29 months (Baldwin et al. 2001). Holton and Tan (1980) found that during the easterly QBO at level 50 hPa the polar vortex was statistically significantly weaker than during westerly QBO at the same level. There is no precise consensus of the mechanisms of this tropical-extratropical connection, but the most common explanation is that the QBO affects the polar vortex via the Holton Tan effect: During easterly QBO, small amplitude planetary waves are reflected back towards the North Pole weakening the polar vortex (Holton and Tan 1980, 1982; Watson and Gray 2014, Gray et al. 2018). Garfinkel et al. (2018) found by model simulation a weakened stratospheric polar vortex during the easterly QBO phase compared to the westerly phase in early winter (October–December).

Challenges related to the realistic modelling of the dynamical stratosphere-troposphere coupling have been adduced by Shepherd et al. (2018) and Polichtchouk et al. (2018). Even though many models are already able to reproduce the QBO, they are still underestimating its teleconnection to the surface weather (Scaife et al. 2014). Furthermore, the anomalous QBO disruption in winter 2015/2016 was not forecasted by the models (Newman et al. 2016).

In this paper, we first verify the raw and the bias-corrected surface temperature reforecasts of the ECMWF's ERFs for forecast weeks 1 to 6 over Northern Europe against ERA-Interim surface temperature re-analysis (Dee et al. 2011). After that, our aim is to find out which stratospheric observations available at the start of the forecast are followed by a statistically significantly weaker AO index. For this, we explore the observed daily AO index during boreal winters 1981-2016, 1–2 weeks, 3–4 weeks, and 5–6 weeks after different phases of QBO and strengths of the observed stratospheric zonal mean winds. According to the observed daily AO index, after different phases of QBO, and strengths of the observed stratospheric zonal mean winds, we define a Stratospheric Wind Indicator (*SWI*), which is a novel indicator for the strength of the AO index in the following 1 to 6 weeks. For a statistically significantly weaker mean AO index, the *SWI* is defined as SWI_{neg}; otherwise, *SWI* is defined as SWI_{plain}. Further, we study the mean surface temperature anomalies observed in Northern Europe 1–2 weeks, 3–4 weeks, and 5–6 weeks after SWI_{neg} versus SWI_{plain} and utilize these in post-processing the mean of the temperature forecasts of ECMWF reforecasts. Finally, we compare the *SWI* based post-processed ECMWF reforecasts with the mean bias-corrected ECMWF reforecasts. Our paper is constructed as follows: First, we present the datasets and methods. Then, we present results about the selection of the SWI_{neg} and SWI_{plain} and the skill scores of the forecasts without post-processing and with post-processing. In the Discussions and Conclusions section, we present our view on our findings and the possible next steps.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

2 Datasets and Methods

We verified and post-processed ERFs of the ECMWF's Integrated Forecasting System (IFS Cycle 43r1, Vitart, 2014), which belongs to the models of the Sub-seasonal to seasonal scale (S2S) prediction project of the World Weather Research Program/World Climate Research Program (Vitart et al. 2017). These forecasts are run twice a week, on Mondays and Thursdays, in a horizontal resolution of 0.4 degrees. We studied the weekly mean temperatures of the Monday runs over Northern Europe in forecast weeks 1 to 6. We verified the 20 years × 52 weeks = 1040 reforecasts (11 members ensemble) for 1997–2016 run for the same dates as the operational forecasts, i.e., as Mondays in 2017. The weekly averages of the raw, mean bias-corrected (Section 2.2), and post-processed (Sections 2.3 and 2.4) surface temperature forecasts over Northern Europe were verified against ERA-Interim 1981–2016 temperature re-analyses (Dee et al. 2011). Years 1981–2010 of the ERA-Interim data were used as the climatological reference period and as the statistical/climatological forecast.

2.1 Skill scores of the forecasts

A commonly used measure for the probabilistic forecasts is the continuous ranked probability score (CRPS, Hersbach 2000) calculated by the following Eq. (1):

$$\boldsymbol{CRPS} = \int |\boldsymbol{F(y)} - \boldsymbol{F_o(y)}|^{\boldsymbol{2}} \boldsymbol{dx}, \quad (1)$$

where $F(y)$ and $F_o(y)$ are the cumulative distribution functions of the forecast and the observation, respectively.

The CRPSs were calculated by the R package 'ScoringRules' (Jordan et al. 2018) for the ECMWF's reforecast ($CRPS_{rf}$) and the climatological forecasts (ERA-Interim weekly mean temperatures in 1981–2010), which were used as the reference ($CRPS_{clim}$). As the ensemble size of the reforecasts, m, was only 11, and the ensemble size of the operational forecasts, M, was 51, the expected CRPS, the $CRPS_{RF}$ of the $CRPS_{rf}$ was calculated for 51 members using equation 26 in Ferro et al. (2008):

$$CRPS_{RF} = \frac{m(M+1)}{M(m+1)} CRPS_{rf} \quad (2).$$

Further, the skill scores of the CPRS, the CRPSS, were calculated as follows:

$$CRPSS = 1 - \frac{CRPS_{RF}}{CRPS_{clim}} \quad (3).$$

The annual mean CRPSSs were calculated for the ECMWF's reforecasts (1997–2016) for forecast weeks 1 to 6. The statistical significances of each forecast week's annual mean CRPSS was determined for each grid point. The p-value with the null hypothesis that the CRPSS is zero was calculated by bootstrap resampling procedure with replacement and a sample size of 5000 for significance level 0.05.

2.2 Bias correction of the ensemble mean

The mean bias correction (as in Buizza and Leutbecher 2015, eq. 7a) removed the mean bias computed from the ensemble reforecasts for the 20 years (1997–2016) depending on the forecast week date. For the 1997–2016 reforecasts, the average bias was calculated considering 19 × 11 × 5 = 1045 ensemble reforecast members: 11 members' reforecast with initial dates defined

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

by five weeks centred on the forecast week date for the 19 years reforecasts (1997–2016 excluding the reforecast year). The mean bias-corrected weekly mean temperatures were verified against the ERA-Interim data by calculating the annual mean CRPS separately for each forecast week (1 to 6). The skill scores of the mean bias-corrected forecasts and their statistical significance were calculated as explained in Section 2.1.

2.3 Definition of the stratospheric wind indicator (*SWI*)

As numerous observational and modelling studies have shown, the stratospheric polar vortex influences the weather in the Northern Hemisphere during boreal winter; strong polar vortex coincides more often with positive AO index and mild surface weather in Northern Europe, whereas weak polar vortex is more often followed by negative AO index and cold air outbreaks (Thompson and Wallace 1998, 2001, Kidston et al. 2015 and references therein). We aimed to find stratospheric precursors for a statistically significantly weaker AO index available at the start of the forecast. The observed daily surface AO indexes were downloaded from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC). We used two stratospheric wind data sets for the precursors of the AO index. The first data were the daily ZMZW at 60°N and 10 hPa during 1981–2016 provided by the National Aeronautics and Space Administration (NASA). The second data were the monthly mean zonal wind components at levels 70 hPa, 50 hPa, 40 hPa, 30 hPa, 20 hPa, 15 hPa, and 10 hPa from the Singapore radio soundings, during 1981–2016, provided by the Free University of Berlin, representing the equatorial stratospheric monthly mean zonal wind components, the QBO (Naujokat 1986).

We examined the observed daily AO indexes during the 1–2 weeks, 3–4 weeks, and 5–6 weeks following different phases of QBO and strengths of the stratospheric winds. Holton and Tan (1980, 1982) demonstrated that the geopotential height at high latitudes was significantly lower during westerly QBO compared to the easterly QBO. Therefore, we explored the daily AO index 1–6 weeks after westerly QBO, the *WQBO*, and easterly QBO, the *EQBO* using the QBO winds at 30 hPa. We also examined the effect of the maximum of the monthly mean zonal wind components of the QBO between 70 hPa and 10hPa during *EQBO*. Moreover, we explored the daily AO index after the daily ZMZW at 60°N and 10 hPa during the last 10 days of the previous month falling below its overall wintertime (November–March 1981–2016) 10th percentile, corresponding a value of 3.8m/s, indicating a weak polar vortex already at the start of the forecast. The statistical significance of the difference between the AO index following two different stratospheric situations, e.g., the *EQBO* and the *WQBO*, was determined using a two-sided Student's t-test with the null hypothesis that there is no difference. We used the most statistically significant predictors for weaker AO indexes observed 1–2 weeks, 3–4 weeks, and 5–6 weeks after these stratospheric situations, to define a *SWI* to be SWI_{neg}; otherwise, it was defined as SWI_{plain} for the beginning of each winter month (November–February) in 1981–2016.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

2.4 Utilizing the stratospheric winds indicator (*SWI*) in forecasting

In this section, we investigated the observed and reforecasted surface temperature anomalies 1–2 weeks, 3–4 weeks, and 5–6 weeks after SWI_{neg} and SWI_{plain} defined in Section 2.3. First, we calculated the observed two-week mean temperature anomalies of the ERA-Interim reanalyses (Dee et al. 2011) of the 1–2 weeks, the 3–4 weeks, and the 5–6 weeks from the beginning of the January, February, November, and December in 1981–2016 in Northern Europe. Subsequently, we divided the observed two-week mean temperature anomalies to sets of anomalies, representing SWI_{neg} and SWI_{plain} according to the previous month's stratospheric wind observations. Thereafter, we determined the statistical significance of the difference between the surface temperatures after SWI_{neg} and SWI_{plain} using a two-sided Student's t-test with the null hypothesis that there is no difference between SWI_{neg} and SWI_{plain}. This same procedure to define the difference between the surface temperatures after SWI_{neg} and SWI_{plain} was used for the ERA-Interim reanalyses for the period 1997–2016 to see how the selection of a shorter period affects the temperature anomalies. Further, the mean surface temperature anomalies 1–2 weeks, 3–4 weeks, and 5–6 weeks after SWI_{neg} and SWI_{plain} in the ECMWF reforecasts run of the first week of November–February 1997–2016 were defined to examine how the model reproduced the anomalies.

For post-processing the ECMWF reforecasts, we calculated TA_{SWIneg} and $TA_{SWIplain}$ representing mean temperature anomalies in November–February after SWI_{neg} and SWI_{plain}, respectively. In order to decrease the effect of the time period used for defining mean anomalies, the TA_{SWIneg} and $TA_{SWIplain}$ were calculated from the ERA-Interim 1981–2016 two-week mean temperature anomalies by dividing thirty five years' data (1981–2016 excluding the reforecast year) to 5-year-long periods. One 5-year-long period was left out, and the mean anomalies representing SWI_{neg} and SWI_{plain} were calculated of the remaining 30 years. This was repeated seven times by changing the 5-year-long period left out. The means, TA_{SWIneg} and $TA_{SWIplain}$ of the achieved seven mean temperature anomalies representing SWI_{neg} and SWI_{plain} were calculated separately for each grid 0.4 × 0.4 grid point over Northern Europe.

For the post-processing of the ECMWF reforecasts, we first defined the *SWI* either SWI_{neg} or SWI_{plain} at the start of the forecast according to previous months' stratospheric wind observations. According to the *SWI,* we added either TA_{SWIneg} or $TA_{SWIplain}$ to the ERA-Interim mean temperature during 1981–2016, corresponding to forecast weeks 1–2, 3–4, and 5–6 to get a SWI_{neg} and SWI_{plain} based mean temperatures, T_{SWIneg} and $T_{SWIplain}$, for weeks 1–2, 3–4, and 5–6, respectively. The T_{SWIneg} and $T_{SWIplain}$ were used in post-processing the ECMWF reforecasts' mean bias-corrected ensemble members, T_{BC}, by calculating a weighted average, T_{SWI_BC}, for SWI_{neg} as follows:

$$T_{SWI\_BC} = (1 - k_{SWI}) * T_{BC} + k_{SWI} * T_{SWIneg} \quad (4)$$

And for SWI_{plain},

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

$$T_{SWI\_BC} = (1 - k_{SWI}) * T_{BC} + k_{SWI} * T_{SWIplain} \quad (5)$$

where T_{SWI_BC} was a post-processed ensemble member. k_{SWI} was the weight of the T_{SWIneg} or $T_{SWIplain}$, which was tested between 0–1 and defined according to the best improvement in the skill scores of the post-processed forecast. By Eq. (4) and Eq. (5), we adjusted each ensemble member with the same weight, and hence, the original spread of the ECMWF reforecasts remained unchanged. The skill scores of the *SWI* based post-processed forecasts, and their statistical significance, were calculated as explained in Section 2.1.

3 Results

3.1 Skill scores of the forecasts

The annual mean of the expected CRPSS and its 95% level of confidence of the raw and the mean bias-corrected (Section 2.2) weekly mean temperature of the ECMWF reforecasts for 1997–2016 are displayed in Figure 1. In grid points where the CRPSS was higher than zero and the confidence level was higher than 95% (dotted areas), the reforecasts were statistically significantly better than just the statistical forecast based on 1981–2010 climatology. Figure 1 illustrates that for forecast weeks 1–6 the mean bias-corrected ERF reforecasts were on average significantly better than climatology. The annual mean CRPSS values show that in forecast weeks 1–3 the CRPSSs are for the most part above 0.1, whereas on in forecast weeks 4–6 they are mostly lower, between 0 and 0.1.

3.2 The stratospheric observations and the thereafter observed AO index and surface temperature

Figure 2 shows boxplots of the observed minimum of the daily AO index 1–2 weeks, 3–4 weeks, and 5–6 weeks after different phases of QBO and restrictions in the strength of the stratospheric winds in 1981–2016. The first box (brown) represents the minimum AO indexes after all the cases in 1981–2016 November–February, i.e., 36 years * 4 months =144 cases. The second and third boxes show the minimum AO indexes after easterly (*EQBO*, blue) and westerly (*WQBO*, pink) QBO at the 30 hPa level, respectively. The p-value written below each boxplot pair indicates the likelihood of such a pair of distributions arising from a random sampling of a single distribution as given by a Student's t-test, i.e., p-values less than 0.05 indicate that the means of the data sets differ significantly at the 95% level of confidence. The median and the mean of the minimum AO indexes 1–2 weeks, 3–4 weeks, and 5–6 weeks after *EQBO* were lower than after *WQBO*. However, this was statistically significant at the 95% confidence level only in weeks 1–2. The *EQBO* (blue) box shows all the cases of *EQBO* with no restriction in the QBO's monthly mean zonal wind components, whereas the fourth, the sixth, and the eighth blueish boxes show the minimum AO indexes after *EQBO* with all the QBO's monthly mean zonal wind components between levels 70…10hPa being below 13 m/s, 10 m/s, and 7 m/s, respectively. Restricting the *EQBO* cases by a maximum of the QBO's

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

monthly mean zonal wind components in levels 70…10hPa decreased the median of the minimum AO during the following 1–2, 3–4, and 5–6 weeks; however, only for the QBO's monthly mean zonal wind components less than 10 m/s, the minimum AO index was still at the 95% confidence level statistically significantly lower than in the rest of the data.

The 10th box (yellow) in Fig. 2 shows the minimum AO indexes after cases the daily ZMZW at 60°N and 10 hPa was below its 10th percentile (4.8m/s) during the 10 last days of the previous month, corresponding to cases with weak polar vortex already at the start of the forecast. The observed minimum AO index was statistically significantly at the 95% confidence level weaker 1–2 weeks and 3–4 weeks after the daily ZMZW at 60°N and 10 hPa, which was been below their overall wintertime 10th percentile (indicating a weak polar vortex). In weeks 5–6, the AO index was also weaker, but it was statistically insignificant at the 95% level.

Aiming to select stratospheric precursors indicating weak AO with the greatest statistical significance, we defined the *SWI* to be negative in cases when the QBO was easterly at 30 hPa and the QBO's monthly mean zonal wind components in levels 70…10hPa were weaker than 10m/s and/or if the daily ZMZW at 60°N and 10 hPa during the 10 last days of the previous month fell below its overall wintertime 10th percentile. In other cases, the *SWI* was defined as plain. This decision tree for the *SWI* is depicted in Fig. 3. The means of the minimum AO index after SWI_{neg} and SWI_{plain} (in 1981–2016) were statistically significantly different using a Student's t-test, with lower AO index more common 1–2 weeks, 3–4 weeks, and 5–6 weeks after SWI_{neg} than after SWI_{plain} (see Fig. 2 for the p-values).

Figure 4 shows the observed (periods 1981–2016 and 1997–2016) and model forecasted (the period 1997–2016) mean temperature anomalies of the weeks' 1–2, 3–4, and 5–6 in November–February after SWI_{neg} and SWI_{plain}. The observations showed on average lower mean temperatures for the weeks' 3–4 and 5–6 after SWI_{neg} (Fig. 4b-c and 4h-i). The reforecasts also showed cold anomalies after SWI_{neg} (Fig. 4n-o) but weaker than the observed ones. Further, the observations showed on average higher mean temperatures for weeks 1–2, 3–4, and 5–6 after SWI_{plain} (Fig. 4d-f and 4j-l). This warm anomaly was forecasted only to some degree in the forecasts weeks 1–2 (Fig. 4p), and it was totally absent in forecast weeks 3–4 (Fig. 4q) and 5–6 (Fig. 4r). The mean temperature anomalies 3-6 weeks after SWI_{neg} (Fig. 4b-c) and SWI_{plain} (4e-f) during 1981–2016 were statistically significantly different using a Student's t-test, with anomalously cold surface temperatures more common 3-6 weeks after SWI_{neg}. When examining the years 1997–2016 (Fig. 4h-i and Fig. 4k-l), which was the reforecast period, the temperature anomalies were of the same sign than during the longer 1981–2016 period (Fig. 4b-c and Fig. 4e-f) but weaker and not statistically significant all over the Northern Europe.

3.3 The *SWI* and the forecasted mean temperatures

The mean temperature anomalies in Fig. 4(a-f) for Northern Europe were used for the *SWI* based post-processing as described in Section 2.4. The CRPSS of the mean temperature of the forecast weeks 1–2 were not improved by the *SWI* (no figure),

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

whereas the CRPSSs of the mean temperatures of the forecast weeks 3–4 and 5–6 were improved by the *SWI* based post-processing (Fig. 5a and 5b). The best median CRPSS was achieved by k_{SWI}=0.5, for both weeks 3–4 and weeks 5–6. Figure 6 shows the forecasts skill of the mean temperature of the forecast weeks 3–4 and weeks 5–6 forecasted by the mean bias-corrected reforecasts alone (Fig. 6a-b) and by the mean bias-corrected reforecasts together with the *SWI* based post-processed forecast (k_{SWI} =0.5, Fig. 6c-d). By using the *SWI* based post-processing to the ECMWF forecasts, the CRPSSs for weeks 3–4 and weeks 5–6 were slightly improved and the area of these forecasts being significantly better than just the climatological forecast was expanded. The forecast skills for the weeks 3–4 and 5–6 post-processed by Eq. (4) and Eq. (5) were not sensitive to the period (within 1981–2016) used for defining SWI_{neg} and SWI_{plain} temperature anomalies. For instance, we tested periods 1981–2000 and 2001–2016 and got almost the same CRPSSs for Northern Europe (no figure).

4 Discussion and Conclusions

Based on ECMWF's extended-range reforecasts for the period 1997–2016, we found that the weekly mean surface temperature forecasts over Northern Europe were on average significantly better than just the climatological forecast in weeks 1–6, however, in weeks 4–6, the CRPSSs were quite low, mostly between 0 and 0.1.

We showed that in addition to the previously demonstrated weaker polar vortex during easterly QBO in comparison to westerly QBO (e.g., Holton and Tan 1980, Garfinkel et al. 2018), the maximum strength of the QBO's monthly mean zonal wind components in levels 70…10hPa during the easterly QBO at 30 hPa affected the observed AO index 1–6 weeks later. Based on observations, we found that the minimum AO index was statistically significantly weaker 1–2 weeks, 3–4 weeks, and 5–6 weeks after the monthly mean QBO was easterly at 30 hPa, and all the QBO's monthly mean zonal wind components in levels 70…10hPa were less than 10 m/s. We also found that the minimum AO index was statistically significantly weaker 1–2 weeks and 3–4 weeks after the daily ZMZWs at 60°N, and 10 hPa had been below their overall wintertime 10th percentile (indicating a weak polar vortex). In weeks 5–6, the AO index was weaker but statistically insignificant.

Selecting the SWI_{neg} to include the both above-mentioned situations, the level of statistical significance of a weaker AO index during the next 1–2, 3–4, and 5–6 weeks decreased in comparison to using only one of these situations. Our definition of SWI_{neg} resulted in a statistically significantly weaker AO index within the following 1-6 weeks in comparison to the rest of the data, defined as SWI_{plain}. Also, the mean surface temperature anomalies in Northern Europe in November–February in 1981–2016 after SWI_{neg} and SWI_{plain} were statistically significantly different, with anomalously cold surface temperatures more common 3–6 weeks after SWI_{neg}. The mean temperature anomalies corresponding SWI_{neg}/SWI_{plain} at the start of the forecast were used in post processing the ECMWF's mean temperature reforecast for weeks 3–4 and 5–6 in Northern Europe during boreal winter, and thereby, those weeks' forecast skills were slightly improved.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

This study demonstrates that the QBO-polar vortex connection should be better integrated into the extended-range surface temperature forecasts over Northern Europe. The *SWI* based post-processing method introduced in this paper could also be tested for other northern areas affected by the polar vortex and to precipitation and windiness forecasts, and it could be further developed by, e.g., the Madden-Julian-Oscillation (Madden and Julian 1994; Zhang 2005; Jiang et al. 2017; Vitart 2017; Vitart and Molteni 2010; Robertson et al. 2018, Cassou 2008). In this study, the effect of global warming was not filtered from the temperature anomalies used for statistical post-processing. In future work, the impact of filtering the effect of global warming could be tested. In our studies, the use of monthly mean stratospheric observations restricted the studies to post processing only the forecasts made on the first week of each month. The next step would be looking for the stratospheric signal from the forecast model, which would also make it possible to post-process forecasts made every week.

Data availability. ERA-Interim data available at https://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/ (last accessed 24 June 2019). ECMWF reforecasts data available at https://apps.ecmwf.int/mars-catalogue/ (last accessed 28 June 2019). AO indexes data available at https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/ao.shtml (last accessed 24 June 2019). The daily ZMZW at 60°N and 10 hPa data available at https://acd-ext.gsfc.nasa.gov/Data_services/met/ann_data.html (last accessed 24 June 2019). The QBO data data available at https://www.geo.fu-berlin.de/met/ag/strat/produkte/qbo/qbo.dat (last accessed 24 June 2019). The data of Figures 1–2 and 4–6 available at https://github.com/fmidev/sixweeks.

Competing interests. The authors declare that they have no conflict of interest.

Author contributions. NK designed the study, analysed the results and prepared the manuscript with contributions from all co-authors. OH participated in the study design and analysing the results. MK contributed to the discussions and fine-tuned the experiments. DSR contributed to the discussions and to the interpretation of the results. HJ provided supervision during the experiments and writing. HG contributed to the study design and was in charge of the management and the acquisition of the financial support for the CLIPS-project leading to this publication.

Acknowledgements. We wish to thank Academy of Finland for funding the project (number 303951 SA CLIPS). We also acknowledge the ECMWF for monthly forecast data and ERA-Interim data, NOAA/CPC for providing the AO index data, NASA for providing 10hPa wind data, and Free University of Berlin for providing the QBO data. We thank the CLIPS team and developers of the R cran calculation package 'ScoringRules'.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

References

Baldwin, M. P., and Dunkerton, T.J.: Propagation of the Arctic Oscillation from the stratosphere to the troposphere, J. Geophys. Res, 104, D24, 30937-30946, https://doi.org/10.1029/1999JD900445, 1999.

Baldwin, M. P., and Dunkerton, T. J.: Stratospheric harbingers of anomalous weather regimes. Science, 294, 581–584, doi:10.1126/science.1063315, 2001.

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., et al.: The quasi-biennial oscillation, Rev. Geophys., 39(2), 179–229, 2001.

Buizza, R. and Leutbecher, M.: The forecast skill horizon, Q. J. R. Meteorol. Soc., 141, 3366–3382, doi:10.1002/qj.2619, 2015.

Butler, A. H., Seidel, D. J., Hardiman, S. C., Butchart, N., Birner, T., and Match, A.: Defining sudden stratospheric warmings, Bull. American Meteor. Soc., 96, 1913-1928, doi: http://dx.doi.org/10.1175/BAMS-D-13-00173.1, 2015.

Cassou C.: Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. Nature, 455, 523–527, 2008.

Chambers, J. M., Cleveland, W. S., Kleiner, B, and Tukey, P.A.: Graphical Methods for Data Analysis, The Wadsworth statistics/probability series. Wadsworth and Brooks/Cole, Pacific Grove, CA, 1983.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553–597, 2011.

Ervasti, T., Gregow, H., Vajda, A., Laurila, T. K., and Mäkelä, A.: Mapping users' expectations regarding extended-range forecasts. Adv. Sci. Res., 15, 99–106, doi: 10.5194/asr-15-99-2018, 2018.

Ferro C. A. T., Richardson, D. S., and Weigel, A. P.: On the effect of ensemble size on the discrete and continuous ranked probability scores. Meteorol. Appl. 15: 19–24, doi: 10.1002/met.45, 2008.

Garfinkel, C. I., Schwartz, C., Domeisen, D. I. P., Son, S-W., Butler, A. H., and White, I. P.: Extratropical stratospheric predictability from the Quasi-Biennial Oscillation in Subseasonal forecast models, J. Geophys. Res: Atmospheres, doi: 10.1029/2018JD028724, 2018.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems. Wea. Forecasting, 15, 559–570, doi:10.1175/1520-0434(2000)015,0559: DOTCRP.2.0.CO;2, 2000.

Gray, L. J., Anstey, J. A., Kawatani, Y., Lu, H., Osprey, S., and Schenzinger, V.: Surface impacts of the Quasi Biennial Oscillation, Atmos. Chem. Phys., 18, 8227–8247, https://doi.org/10.5194/acp-18-8227-2018, 2018.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Holton, J. R. and Tan, H. C.: The influence of the equatorial quasi-biennial oscillation on the global circulation at 50mb, J. Atmos. Sci., 37, 2200–2208, 1980.

Holton, J. R. and Tan, H. C.: The quasi-biennial oscillation in the Northern Hemisphere lower stratosphere, J. Meteor. Soc. Japan, 60, 140–148, 1982.

Kidston, J.; Scaife, A. A.; Hardiman, S. C.; Mitchell, D. M.; Butchart, N.; Baldwin, M. P., and Gray, L. J.: Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. Nature Geoscience, 8(6), 433–440, 2015.

Limpasuvan, V., Hartmann, D. L., Thompson, D. W. J., Jeev, K., and Yung, Y. L.: Stratosphere-troposphere evolution during polar vortex intensification. J. Geophys. Res., 110, D24101, doi: 10.1029/2005JD006302, 2005.

Madden, R. A., and Julian, P. R.: Observations of the 40–50-day tropical oscillation—A review. Mon. Wea. Rev., 122, 814–837, 1994.

Jiang, Z., Feldstein, S. B., and Lee S.: The relationship between the Madden–Julian Oscillation and the North Atlantic Oscillation. Q. J. R. Meteorol.Soc. 143: 240–250, January 2017 A DOI:10.1002/qj.2917, 2017.

Monhart, S., Spirig, C., Bhend, J., Bogner, K., Schär, C., and Liniger, M. A.: Skill of subseasonal forecasts in Europe: Effect of bias correction and downscaling using surface observations. J. Geophys. Res: Atmospheres, 123, 7999–8016. 2018.

Jordan A., Krueger F., and Lerch, S.: Evaluating Probabilistic Forecasts with scoringRules. Journal of Statistical Software, forthcoming, 2018.

Naujokat, B.: An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. J. Atmos. Sci., 43, 1873–1877. 1986.

Newman, P. A., L. Coy, S. Pawson, and L. R. Lait: The anomalous change in the QBO in 2015–2016, Geophys. Res. Lett., 43, 8791–8797, 2016.

Polichtchouk, I., Shepherd, T. G., Byrne, N. J.: Impact of Parametrized Nonorographic Gravity Wave Drag on Stratosphere-Troposphere Coupling in the Northern and Southern Hemispheres, Geophys. Res. Lett., 45, 8612-8618, doi: 10.1029/2018gl078981, 2018.

Robertson, A. W., Camargo, S. J., Sobel, A., Vitart, F., and Wang, S.: Summary of workshop on sub-seasonal to seasonal predictability of extreme weather and climate. npj Climate and Atmospheric Science, 1, 8, doi: 10.1038/s41612-017-0009-1, 2018.

Scaife, A. A., et al.: Predictability of the quasi-biennial oscillation and its northern winter teleconnection on seasonal to decadal timescales, Geophys. Res. Lett., 41, 1752–1758, doi:10.1002/2013GL059160, 2014.

Shepherd T. G., Polichtchouk, I., Hogan, R., Simmons, A. J.: Report on Stratosphere Task Force, ECMWF Technical Memorandum n. 824, doi: 10.21957/0vkp0t1xx, 2018.

Thompson, D. W. J. and Wallace, J. M.: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. Geophys. Res. Lett., 25, 1297– 1301, 1998.

Schoeberl, M. R.: Stratospheric warmings: Observations and theory. Rev. Geophys., 16, 521–538, 1978.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

Thompson, D. W. J., Baldwin, M. P. and Wallace J. M.: Stratospheric connection to Northern Hemisphere wintertime weather: implications for prediction. J. Clim. 15, 1421–1428, 2002.

Thompson, D. W. J., and Wallace, J. M.: Regional Climate Impacts of the Northern Hemisphere Annular Mode. Science, 293, 85–89, 2001.

Tomassini, L., Gerber, E. P., Baldwin, M. P., Bunzel, F. and Giorgetta, M.: The role of stratosphere troposphere coupling in the occurrence of extreme winter cold spells over northern Europe. J. Adv. Model. Earth Syst., 4, M00A03, 2012.

Vitart F., and Molteni F.: Simulation of the MJO and its teleconnections in the ECMWF forecast system. Q. J. R. Meteorol. Soc., 136, 842–855, 2010.

Vitart F.: Evolution of ECMWF sub-seasonal forecast skill scores. Q. J. R. Meteorol. Soc., 140, 1889–1899, doi: 10.1002/qj.2256, 2014.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., Hendon, H., Hodgson, J., Kang, H., Kumar, A., Lin, H., Liu, G., Liu, X., Malguzzi, P., Mallas, I., Manoussakis, M., Mastrangelo, D., MacLachlan, C., McLean, P., Minami, A., Mladek, R., Nakazawa, T., Najm, S., Nie, Y., Rixen, M., Robertson, A. W., Ruti, P., Sun, C., Takaya, Y., Tolstykh, M., Venuti, F., Waliser, D., Woolnough, S., Wu, T., Won, D., Xiao, H., Zaripov, R., and Zhang L.: The Subseasonal to Seasonal (S2S) Prediction Project Database. Bull. Amer. Meteor. Soc., 98, 163–173, https://doi.org/10.1175/BAMS-D-16-0017.1, 2017.

Vitart, F.: Madden-Julian Oscillation prediction and teleconnections in the S2S database: MJO prediction and teleconnections in the S2S database. Q. J. R. Meteor. Soc., 143, 2210–2220, 2017.

Watson, P. A., and L. J. Gray: How Does the Quasi-Biennial Oscillation Affect the Stratospheric Polar Vortex? J. Atmos. Sci.,71, 391–409, doi: 10.1175/JAS-D-13-096.1, 2014.

Zhang, C.: Madden-Julian Oscillation. Rev. Geophys., 43, RG2003, doi:10.1029/2004RG000158, 2005.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

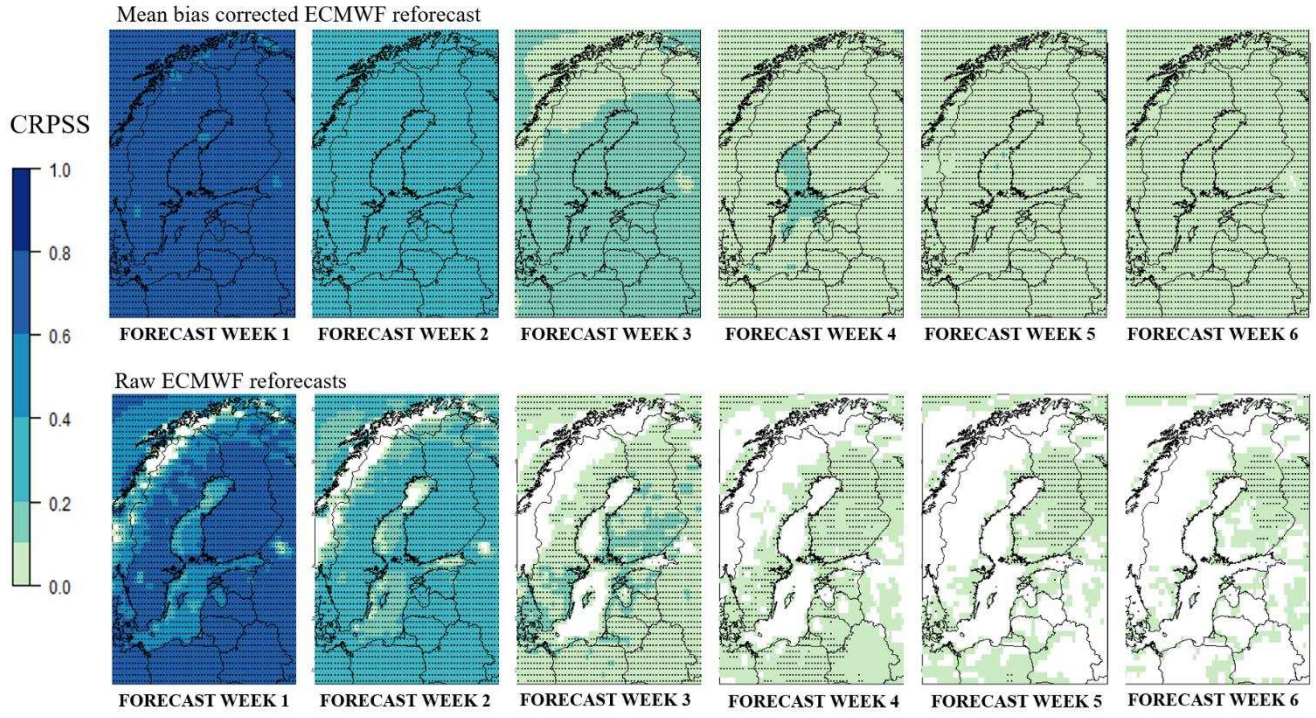

Figure 1: Expected CRPSS of the weekly mean temperature of the mean bias-corrected (upper row) and raw (lower row) ECMWF reforecasts for years 1997–2016 using ERA-Interim climatology of 1981–2010 as the reference. The dotted areas represent the 95% level of confidence that the CRPSS is above zero.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

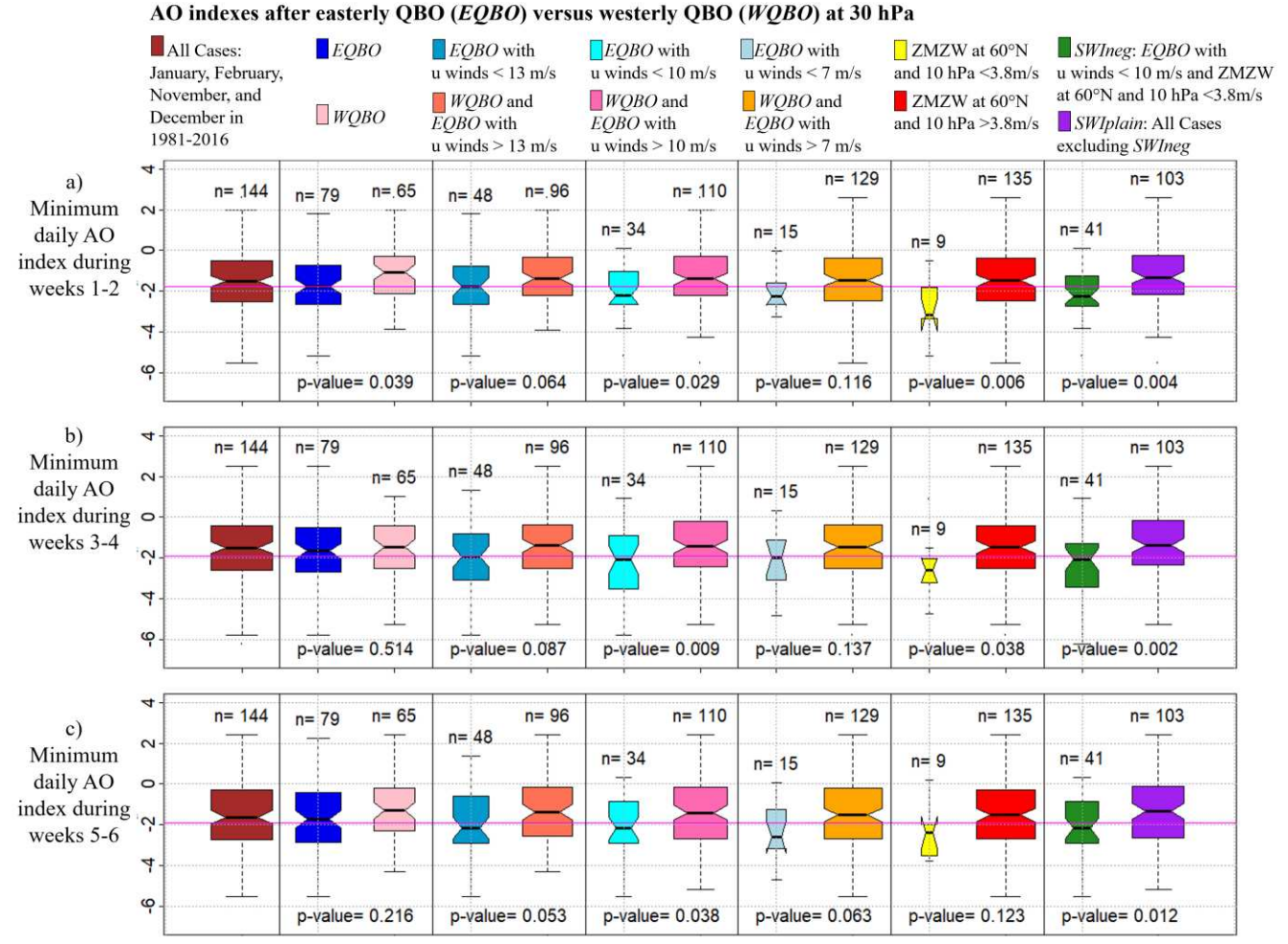

Figure 2: Observed minimum daily AO index a) 1–2, b) 3–4 and c) 5–6 weeks after different stratospheric situations. The line dividing each box into two parts shows the median of the data, the ends of the box show the lower and upper quartiles, and the whiskers represent the highest and the lowest values excluding outliers. The *n* written above each box indicates the number of observations in each group. The widths of the boxes have been drawn proportional to the square-roots of *n*. The p-value written below each boxplot pair indicates the likelihood of such a pair of distributions arising from a random sampling of a single distribution as given by a Student's t-test, i.e., p-values less than 0.05 indicate that the means of the data sets differ significantly at the 95% level of confidence. The notches of each side of the boxes were calculated by R boxplot.stats. If the notches of two plots do not overlap, this is 'strong evidence' that the two medians differ (Chambers et al., 1983, p. 62). The magenta line represents the mean minus one standard deviation of the daily AO index of all the cases. ZMZW=zonal mean zonal wind.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

Atmospheric Chemistry and Physics Discussions

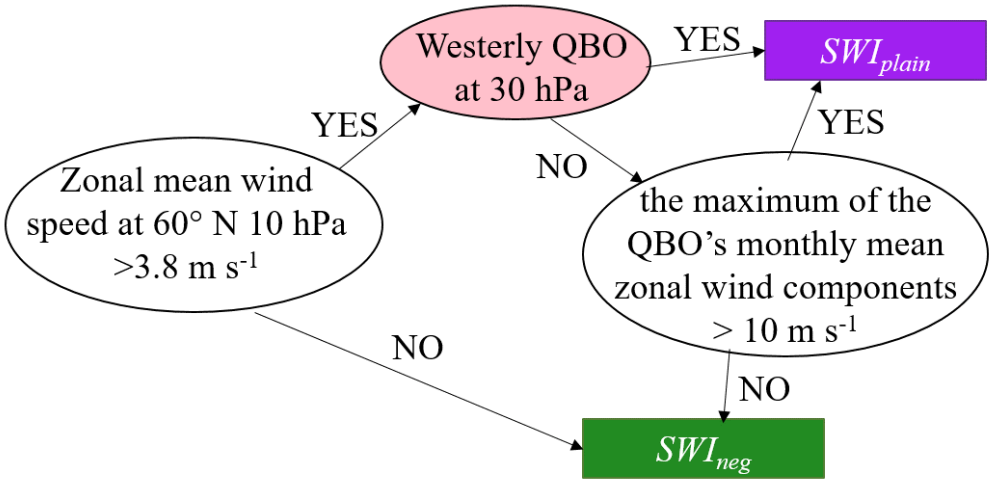

Figure 3: Decision tree of SWI_{neg}/SWI_{plain}.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

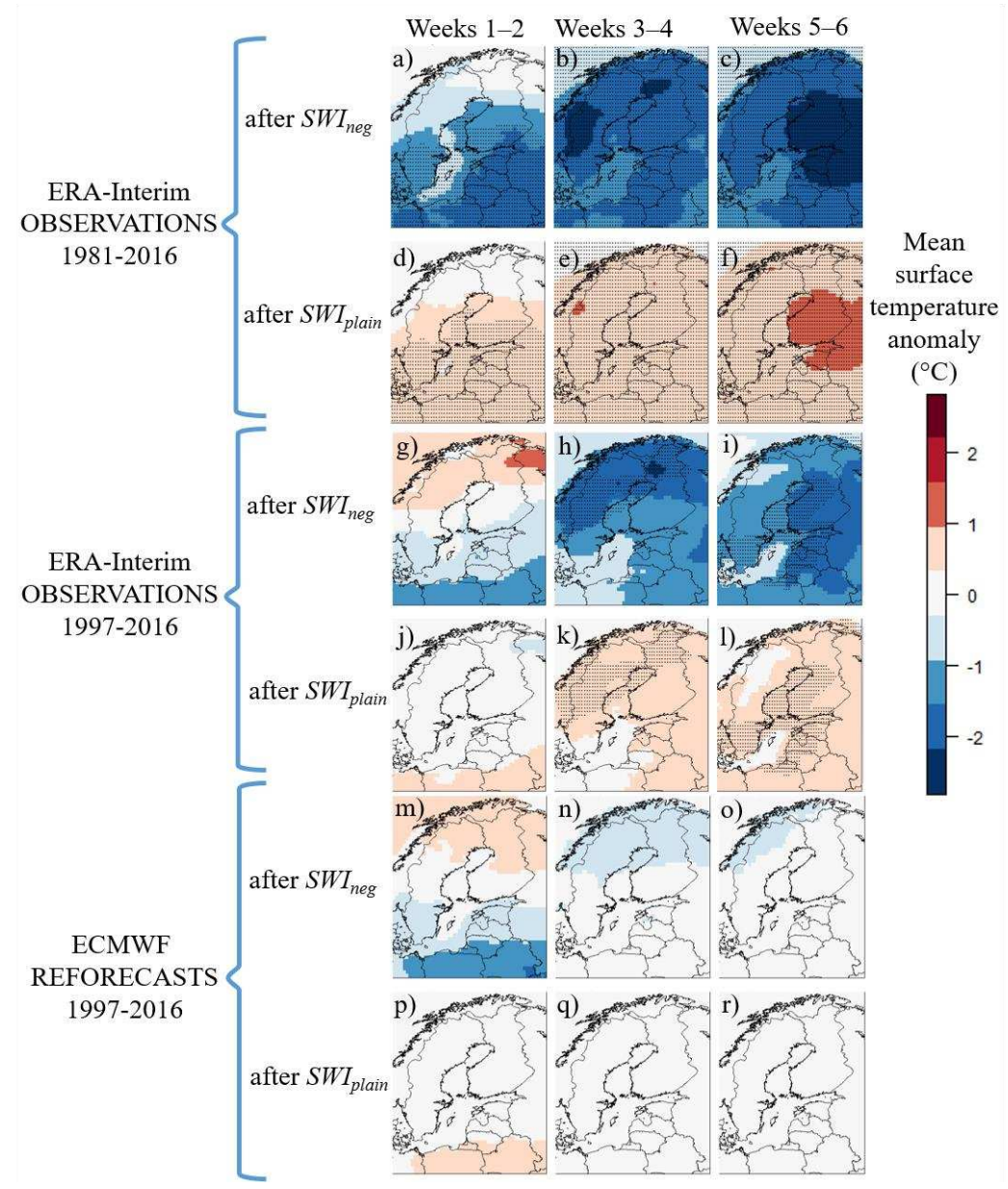

Figure 4. ERA-Interim observed (a-l) and ECMWF reforecasted (m-r) mean temperature anomalies during boreal winters (November-February) in cases the previous month's *SWI* was negative (SWI_{neg}, covering about 28% of the winter months) or plain (SWI_{plain}, covering about 72% of the winter months). The dotted areas represent the 95% level of confidence where the means of surface temperature anomalies after SWI_{neg} and SWI_{plain} differ significantly.

https://doi.org/10.5194/acp-2019-679
Preprint. Discussion started: 19 August 2019
© Author(s) 2019. CC BY 4.0 License.

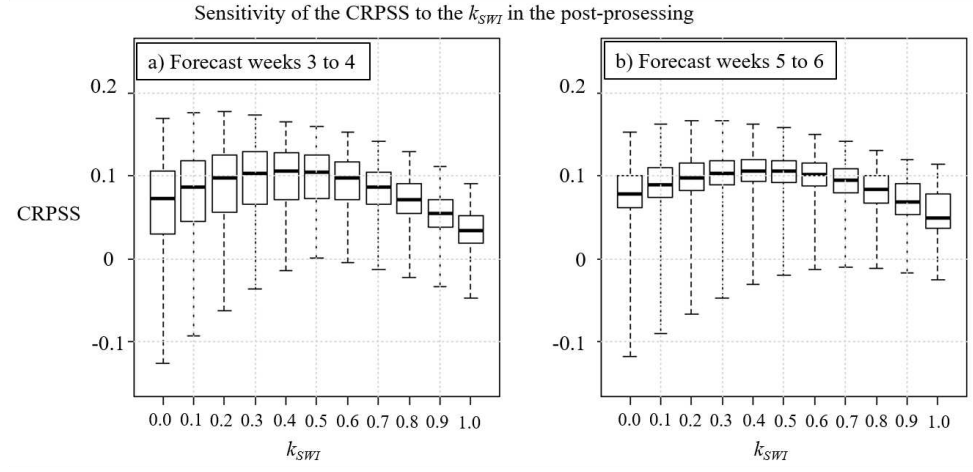

Figure 5. Sensitivity of the expected CRPSS of the ECMWF surface temperature reforecasts to the k_{SWI} ranging from 0.0 to 1.0 in forecast weeks 3–4 (a) and 5–6 (b). The black boxes show the lower and upper quartiles, and the whiskers illustrate the extremes of the November-February mean CRPSSs of all the grid points in Northern Europe.

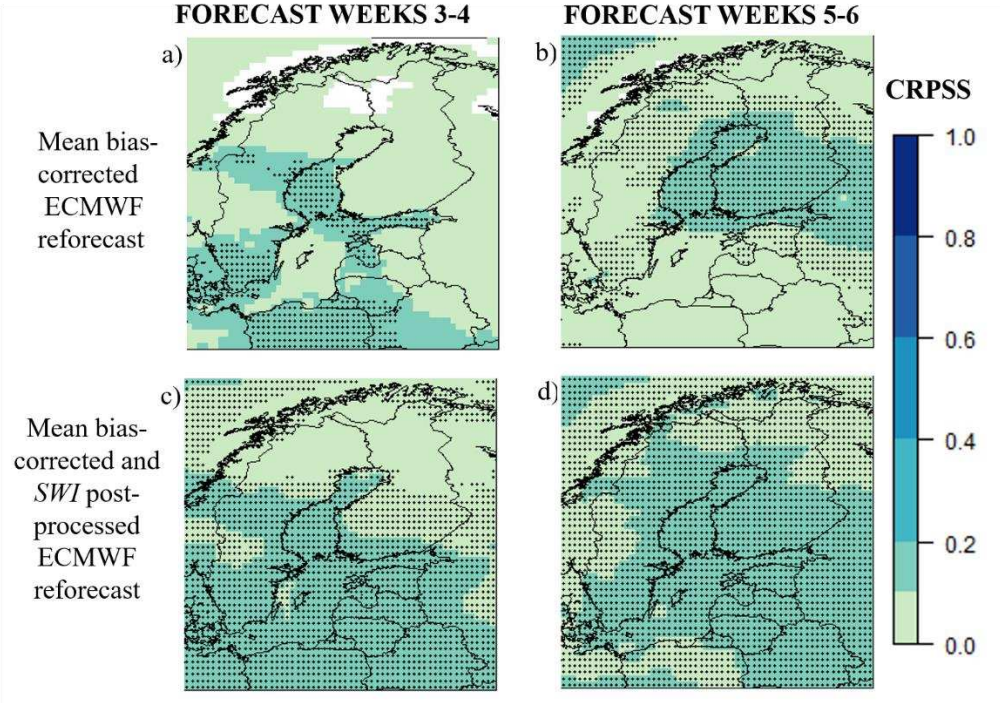

Figure 6. Expected CRPSS of forecast weeks 3–4 and 5–6 of the ECMWF's mean temperature reforecasts for November–February 1997–2016 after mean bias-correction (a-b) and after both mean bias-correction and the *SWI* based post-processing (c-d). ERA-Interim climatology of 1981–2010 was used as the reference. The dotted areas represent the 95% level of confidence that the CRPSS is above zero.

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

FINNISH METEOROLOGICAL INSTITUTE
Erik Palménin aukio 1
P.O. Box 503
FI-00560 HELSINKI
tel. +358 29 539 1000
WWW.FMI.FI

FINNISH METEOROLOGICAL INSTITUTE
CONTRIBUTIONS No. 161
ISSN 0782-6117
ISBN 978-952-336-096-9 (paperback)
ISBN 978-952-336-097-6 (pdf)
https://doi.org/10.35614/isbn.9789523360976
Helsinki, 2020

Edita Prima Oy